

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	901	41.8		363	13	Q93413	Q93413 spheroideae
	2	845	39.2		374	13	Q93412	Q93412 spheroideae
	3	602	27.9		559	13	Q93414	Q93414 spheroideae
	4	484.5	22.5		426	4	Q9HB89	Q9HB89 homo sapien
	5	475.5	22.1		395	11	Q91276	Q91276 mus musculu
	6	475	22.0		415	4	Q96AM5	Q96AM5 homo sapien
	7	474	22.0		415	4	Q96Z04	Q96Z04 homo sapien
	8	472.5	21.9		395	11	Q9JUT1	Q9JUT1 ratius norv
	9	470.5	21.8		395	11	Q9ESQ4	Q9ESQ4 ratius norv
	10	456.5	21.2		405	11	O55040	O55040 mus musculu
	11	434	20.1		418	5	O17239	O17239 caenorhabdi
	12	430	19.9		428	5	Q9JFN4	Q9JFN4 drosophila
	13	414	19.2		595	5	Q9VFW6	Q9VFW6 drosophila
	14	410	19.0		402	11	Q9JTR2	Q9JTR2 ratius norv
	15	405	18.8		410	4	Q9BPH6	Q9BPH6 homo sapien
	16	405	18.8		412	11	Q9JUT5	Q9JUT5 ratius norv

17	396.5	18.4	416	11	Q8VFE5	Q8vff5 mus musculus
18	395.5	18.3	416	11	Q920O5	Q920g5 mus musculus
19	386	17.9	660	5	Q9VEF5	Q9VEF5 drosophila
20	378	17.5	378	5	Q187O1	Q187o1 caenorhabdi
21	369.5	17.1	385	11	Q9JK40	Q9JK40 mus musculus
22	353.5	16.4	401	13	Q9DDP0	Q9ddo xenopus lae
23	340.5	15.8	370	13	Q8WVL5	Q8wvl5 fugu rubrip
24	339	15.7	145	6	Q97914	Q97914 ovis aries
25	338.5	15.7	390	13	Q8OC04	Q8ocg4 carassius a
26	338	15.7	352	11	Q88820	Q88820 rattus norv
27	338	15.7	352	11	Q9R297	Q9r297 rattus norv
28	338	15.7	397	13	Q9DDP1	Q9ddr1 xenopus lae
29	337.5	15.7	513	13	Q9DFP9	Q9dfp9 catostomus
30	335	15.5	352	11	Q9OWM3	Q9owm3 rattus norv
31	334	15.5	382	11	Q9ERT2	Q9ert2 mus musculu
32	329	15.3	404	13	Q9DPB0	Q9dpb0 catostomus
33	327.5	15.2	379	5	Q8T760	Q8t760 brachiostom
34	326	15.1	521	5	Q9VAD2	Q9vad2 drosophila
35	325	15.1	359	13	Q9PEV7	Q9pev7 anguilla an
36	325	15.1	335	5	Q9SYD7	Q9syd7 caenorhabdi
37	323	15.0	436	13	Q73868	Q73868 coturnix co
38	320.5	14.9	731	5	Q9VUC4	Q9vdc4 drosophila
39	320	14.8	380	13	Q9DG06	Q9dgg6 carassius a
40	320	14.8	517	5	Q8TOS8	Q8tos8 drosophila
41	319.5	14.8	338	4	Q9UDN6	Q9udn6 homo sapien
42	316	14.7	352	11	Q9JIN4	Q9jin4 rattus norv
43	316	14.7	352	11	Q8VIT4	Q8v7t4 mus musculu
44	315	14.6	377	13	Q98U14	Q98u14 brachiandio
45	314.5	14.6	400	6	Q95M54	Q95m54 macaca fasc

ALIGNMENTS

RESULT 1

ID	0093413;	PRELIMINARY;	PRT;	363 AA.
AC	0093413;			
DT	01-NOV-1998 (TrEMBLrel. 08, Created)			
DT	01-NOV-1998 (TrEMBLrel. 08, last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, last annotation update)			
DE	Orphan G protein-coupled receptor.			
OS	Spherooides nephelus			
OC	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodoniformes;			
OC	Tetraodontidae; Spherooides.			
OX	NCBI_TaxID=39110;			
RN	1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20092336; PubMed=10628755;			
RA	Palyia O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,			
RA	Gao Y.D., Schlaiss K.D., Yang L., Morfello G.J., Nargund R.,			
RA	Patchett A.A., Howard A.D., Smith R.G.;			
RT	Ligand activation domain of human orphan growth hormone (GH)			
RT	secretagogue receptor (GHS-R) conserved from pufferfish to humans.			
RL	Mol. Endocrinol. 14:160-169(2000).			
EMBL	AF082210; AAC33473.1; -			
DR	Interpro: IPRO00276; GPCr_Rhodopsin.			
DR	Pfam: PF00001; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCRHHOOPS.			
DR	PROSITE; PR01565; NEUROMEDINTR.			
DR	PROSITE; PS00237; G PROTEIN RECP F1.1; UNKNOWN_1.			
DR	PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.			
KW	Receptor.			
SO	SEQUENCE	363 AA;	41324 MW;	74518BD9CD310991 CRC64;

```

Query Match      41.8%; Score 901; DB 13; Length 363;
Best Local Similarity 50.38; Pctd. No. 6,7e-68;
Matches 179; Conservative 56; Mismatches 77; Indels 44; Gaps 5;

QY    31 SPFLGLAVLTAVACCLFFVGVGSGNVVTWVLIGRKYRDRTTNLNYLGSMASDILLIG 90
      ||| :||| :: :|||||::| :||| :|||:|||||:|
```

Db 31 SLEPASTLIPVTVICILIFVVGVTGNTMTILLIQYFKDKMTNTNLYSSMAVSDVIFLC 90
 QY 91 LPFDYRLMRSPWVPGPLCRSLVYVEGCTVATLLMTALSVERYLAI CRPLRARVLV 150
 Db 91 LPFDYRLMRSPWVPGPLCRSLVYVEGCTVATLLMTALSVERYLAI CRPLRARVLV 150
 QY 151 TTRRVVALIIVMAVALLSAGPFLVGVGEODPGISVPGVNGTARJASSPLASSPPLML 210
 Db 151 TTRRVVALIIVMAVALLSAGPFLVGVGEODPGISVPGVNGTARJASSPLASSPPLML 210
 QY 151 TTRRVVALIIVMAVALLSAGPFLVGVGEODPGISVPGVNGTARJASSPLASSPPLML 210
 Db 151 TTRRVVALIIVMAVALLSAGPFLVGVGEODPGISVPGVNGTARJASSPLASSPPLML 210
 QY 211 SRAAPPSPSPGPTAAALFSSRECRSPPAQL--GALRVMLWTATVAFPLPCLSLIYNG 268
 Db 186 ---HSDYNTG-----QCKHTGYAISGQLHIMWSTTYFFCPMLCLLFLYG 229
 QY 269 LIGRELMSSRRPLRCGPAASGRGHRQTRKVLVVLVLAFLIICMLPFHVGRITTYNTEDSR 328
 Db 230 SIOGCKMKSKNDIGFCALARESHRQTVILVVLVLAFLIICMLPFHVGRITTYNTEDSR 289
 QY 329 MMVFSQYFNIVALQLFYLSASINPILYNLISKYRAAFAKLLARKSRPRGFRHSR 384
 Db 290 TMLISQNFMAVSLCYLSASINPVLVYVNLMSRKYRAAFAKLLARKSRPRGFRHSR 344

RESULT 2
 ID 093412 PRELIMINARY; PRT; 374 AA.

AC 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Orphan G protein-coupled receptor.
 OS Spherozooids nephelus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Spherozooids.
 NC NCB1_TaxID=39110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20092336; Pubmed=10628755;
 RA Palyna O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
 Rao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
 Patchett A.A., Howard A.D., Smith R.G.;
 RT "Ligand activation domain of human orphan growth hormone (GH)
 RT secretagogue receptor (GHS-R) conserved from pufferfish to humans."
 RL Mol. Endocrinol. 14:160-169(2000).
 DR EMBL: AF082209; AAC33472.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR InterPro: IPR000524; HTH_Gntr.
 DR Pfam: PF00001; 7tm.1.1.
 DR PRINTS: PR00237; GPCRHOPOSN.
 DR PRINTS: PR01565; NEUROMEDINUR.
 DR PRINTS: PR01566; NEUROMEDINUR.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
 DR PROSITE: PS00043; HTH_GNTR_FAMILY; UNKNOWN_1.
 DR KW Receptor.
 SQ SEQUENCE 374 AA; 42324 MW; 2CF9304FD004C7A16 CRC64;

Query Match 39.2%; Score 845; DB 13; Length 374;
 Best Local Similarity 45.5%; Pred. No. 3.6e-63;
 Matches 187; Conservative 64; Mismatches 102; Indels 58; Gaps 12;

QY 5 WNGS--DGEAGAREPPWALPPCDERRCSPFLGALVPTAVACLIFVVGSGNVVTMLI 63
 Db 13 WEGSHNGTAGLELP-----LNYSIPLAVITVACTVLPVAVGAGVMTILVV 61
 QY 64 GRVDMRTTNNLYGSMVSDLLILGLPFDLYRLMRSPWVPGPLCRSLVYVEGCTY 123
 Db 62 SRYRMRRTTNNLYGSMVSDLLIFVCMPLDLYRMRKYPWRFGDLCKLFGVSSCTY 121
 QY 124 ATLLMTALSVERYLAI CRPLRARVLVTRRRVVALIIVMAVALLSAGPFLVGVGEODP 183

Db 122 STILICTALSVERYLAI CRPLARAKALVTKRRVALLILMTVSLISAGPVFMVGEKDS 181
 QY 184 GISVVPGLNGTARJASSPL--ASSPPLMLSRPSPSPGPTAAEALFSSRECRPS--PA 240
 Db 182 --IMFPN-----SSDLNESSMPL-----EADTRECNRMTQAV 212
 QY 241 QLGALRVMLWTATVAFPLPCLSLIYGLIGRELMSSRRPLR--GPAASGRGHRQTRK 299
 Db 213 ESGLEAMVNLSSVFFPMVPCVTLVGLIGRLMLRRETTINGRVAVRDKSNRQTIKM 272
 QY 300 LVLVLAFLIICMLPFHVGRITTYNTEDS--RMVFSQYFNIVALQLFYLSASINPILYN 356
 Db 273 LVLVLAFLIICMLPFHVGRITTYNTEDS--RMVFSQYFNIVALQLFYLSASINPILYN 332
 QY 357 LISKRYRAAFAKLLARKSRPRGFRHSRDTAGEVAGTGGDVGTYETSAN 407
 Db 333 TMSWKYRGAVARLFGVSDSPQ--RGR-TASTVMD-----GMTSTIVS 373

RESULT 3
 ID 093414 PRELIMINARY; PRT; 559 AA.

AC 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Orphan G protein-coupled receptor.
 OS Spherozooids nephelus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Spherozooids.
 NC NCB1_TaxID=39110;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20092336; Pubmed=10628755;
 RA Palyna O.C., Feighner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
 Rao Y.D., Schleim K.D., Yang L., Morriello G.J., Nargund R.,
 Patchett A.A., Howard A.D., Smith R.G.;
 RT "Ligand activation domain of human orphan growth hormone (GH)
 RT secretagogue receptor (GHS-R) conserved from pufferfish to humans."
 RL Mol. Endocrinol. 14:160-169(2000).
 DR EMBL: AF082211; AAC33474.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1.1.
 DR PRINTS: PR00237; GPCRHOPOSN.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
 DR KW Receptor.
 SQ SEQUENCE 559 AA; 61956 MW; AFB9F35FAF0CD5F6 CRC64;

Query Match 27.9%; Score 602; DB 13; Length 559;
 Best Local Similarity 29.5%; Pred. No. 1.6e-42;
 Matches 155; Conservative 65; Mismatches 125; Indels 180; Gaps 12;

QY 22 LPDCERRC---SPFLGALVPTAVACLIFVVGSGNVVTMLIGRYRDMRTTNNLYLG 78
 Db 20 LHKCSNOGCHWEBPV-FGMIIVCVTIIVPLMLFGLIGNITLIIIVVWMLRPRYRSTYLYLS 78
 QY 79 SNAVSDLLILGLPFDLYRLMRSPWVPGPLCRSLVYVEGCTVATLLMTALSVERYL 138
 Db 79 SNAVSDLLILGLPFDLYRLMRSPWVPGPLCRSLVYVEGCTVATLLMTALSVERYL 138
 QY 139 AICRPLRARVLVTRRRVVALIIVMAVALLSAGPFLVGE---ODPGIS----- 186
 Db 139 AVCMPTAKTVVTRRRTRITIGICIMGAISAPVVMVGVGEVQSDOGISGWRSGAW 198
 QY 187 -----VYVGL----- 191
 Db 199 TGKEGKGFIIIGRERENDKGLKDBQLEBMMWKKXEMNCGDKNGVTKGKGDKSLIV 258
 QY 192 -NGTARJASSPLASSPPLMLSRPSPSPGPTAAEALFSSRECRPS--PAQLGA 244

[illegible]

Dd		61	MPCATYLLIFVVGAVNGTLCVLINHKAMRFPNTYYFLSLAVSDLLVLLVGLPLEUDE	120
Oy		98	LMSRPVWFPBPLCRSLVYGEGCTVATLHMALSTERVLACRPLRARLVTRRRYRA	157
Dd		121	MHNHPFLLGVGGCYFTTLIFENVCCLASVNLVALSYERYAAVHPIQASMTVAHYR	180
Oy		158	LIANVAVALLSAGPFLFVEGDEPDGISVPVGINGTARIASSPLMISRAPPSS	217
Dd		181	VLAAGVMGLAMTCLSPNTSLHGIRQ--LNH-----	207
Oy		218	PBGCPFEAALAALFSRECRSPAOLGALRMVLTTR--VFLEPCLSLIYGLIGELW	275
Dd		208	PCRGPVPDASVCMVLP-----RALVMVVQTALLFFCLPMALMSVLYLIGRLR	259
Oy		276	SSRRPL-----RGPAAS-----GREGRGRTKRVLLVVLAFICMLPHVGRIT	320
Dd		260	REHLLMGEAKKGSAABRYTCRIQHDKRGROYTKMLFVLVVFGLICMAPFHADVM	319
Oy		321	Y--INTEDSRMYFSQYFNIVALQLFYLSAINPILYNILSKYRAAPFL-----LARK	374
Dd		320	WSVVSOWTDGLHIAFHVHVHSIGIFFYLGSANPVLSIMSRRFRFGALCIACCHR	379
Oy		375	SRRP--GFHRSDPTAGEVAGDTG	395
Dd		380	LRPRSHSLSRMTTGSTLCIDVG	402
RESULT 5				
O91Z76				
ID		PRELIMINARY:	PRT:	395 AA.
AC		O91Z76:		
DT		01-DEC-2001 (TREMBLrel. 19, Created)		
DT		01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT		01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DS		Neuromedin U receptor type 2.		
OS		Mus musculus (Mouse).		
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.		
OX		NCB1_Taxid=10090;		
RN		[1]		
RP		SEQUENCE FROM N.A.		
RC		STRAIN=C57BL/6;		
RA		Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monsma F.J. Jr.,		
RA		Gustafson E.L.;		
RT		"Characterization of murine neuromedin U R2 receptor.";		
RL		Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.		
DR		EMBL; AY057384; AAL26695.1; -		
DR		InterPro: IPR000276; GPCR_Rhodopsn.		
DR		Pfam: PF00001; Tm 1; 1.		
DR		PRINTS; PRO1565; NEUROMEDINUR.		
DR		PRINTS; PRO1567; NEUROMEDINUR.		
DR		PRINTS; PRO1570; NEPRECEPTOR.		
DR		PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.		
DR		PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.		
KW		Receptor.		
SQ		SEQUENCE	395 AA; 44844 MW; 0BB540024566903D CRC64;	
Query Match		22.1%; Score 475.5; DB 11; Length 395;		
Best Local Similarity		30.6%; Pred. No. 5e-32;		
Matches 115; Conservative 79; Mismatches 101; Indels 81; Gaps 11;				
Oy		30	CSPFPLGALVTVTVACLFLFVVGSGNVVTWLIGRYDMRTTNLYLGSMVAPSDLI-L	88
Dd		29	CGPKRSDLSPVSVYVALIFVVGVIIGNLLVCLVIARHOTIKTPINYYLFSLAVSDLLVLL	88
Oy		89	LGLEPFDLVLRMRSRPWVGPIRLCRSLVYGECCTVATLHMALSYERYYLAICPELARV	148
Dd		89	LGMPELVELEMHNPYLEPGVGCYEKTALPETVCFASTISYTIVSIERYVAIVHFPRAKL	148
Oy		149	LVTRRRVALAILAVMAVALSAGCFLEFIYGV--EQDDPISVPPGINGARTARIASSPLASP	206
Dd		149	ESTRRARLRILISTVMSFSVPSLPTNSIHGIKKPODFPGSSVPG-----SATCYTK	200

```

Oy 207 PLMTSRAPPSPSPCPETNAEAAALFRRRCRSPQAQGLARLVMWLTAYVFF-LPFLCLST 265
Db 201 PMMV-----WSRRPLRGPAASGRGHRQRKVLVVVLAFLICMLP 313
Oy 266 LYGILGREI-----YNTIIQTSLTFYLIMTLISV 226
Db 227 LYLMGLAKLKDESLADKXVTYNHRSR-----KSVMFLFVLVLAICMTP 275
Oy 314 FHVGSIIVINTEDSNMYTFSQYFN--IVALQLFTLSASINPILYNLISKRYRAAPKL 370
Db 276 FHVDR-LFPSPFDEWTESLAAVFNDLIHAVSGVFYLLSAVNPFITYNLISRFR-AAFRNV 333
Oy 371 LA-----RKSRPRG 379
Db 334 VSPSCKWCHPHRPOG 349

RESULT 6
O96AMS ID O96AMS PRELIMINARY; PRT; 415 AA.
AC 096AMS;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neuromedin u receptor 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RL Strausberg R.;
RA Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC016938; AAH16938.1; -
DR InterPro; IPR000276; GPCK_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO1565; NEUROMEDINUR.
DR PRINTS; PRO1567; NEUROMEDNUZ.
DR PRINTS; PRO1570; NPFRCEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KM Receptor.
SQ SEQUENCE 415 AA; 47770 MW; 30BFEDD706436AB9 CRC64;

Query Match 22.0%; Score 475; DB 4; Length 415;
Best Local Similarity 30.8%; Pred. No. 5.8e-32;
Matches 115; Conservative 78; Mismatches 98; Indels 82; Gaps 12

Oy 30 CSPPELLAVPVTAVCLEFVNGSVGNVVTMLIGRYDRNRTTNLYLSMAVSILL-L 88
Db 37 CGRRSHFPLPSVVVYPIFVGVIGLVCLVLIQHQAOKTPTNYLFLSLAVSDLLVL 96
Oy 89 LGDFPDLYRLMRSPWVRGPILCLSIVYEGCVATALLMHTLSVERYAICRPLRARV 148
Db 97 LGMLPEVYEMWRNPFLFGPGCYCFKTALPETVCFASILITIVSERVAAILHPRAKL 156
Oy 149 LVTRRRRALIAVMAVALLSAGPFLVGYE--QDPGISVVPGLNGTARIASSPLASP 206
Db 157 QSTRRALRIILGIWGSVLFSLPNTSHGKHFIYFNGSLVPS-----SATFVIK 208
Oy 207 PLMTSRAPPSPSPGPETAAMAAALFSRECRSPAQQLALRVMLWTTTAYVFF-LPFLCLST 265
Db 209 PMMV-----YNPIIVTSPFLFYLLPMVTIVS 234
Oy 266 LYGILGREIMSR-----RPLRGPAASGRGHRQRKVLVVVLAFLICMLP 313
Db 235 LYLMALRLKKDKSLADEGNANIQRPCR-----KSVNMKLFVLVLAICMAP 283
Oy 314 FHVGSIIVINTED-SRMVYFSQYFIVAL--OLFYLSASINPILYNLISKRYRAAPKL 369
Db 284 FHIDRLFSPVEENSESL-AAVFVLVAVGVSFYYLSSAVNPITLYNLISRRFG-AAFON 340

```

OY 370 ILIARSRPGRFHR 382
 ::||:
 Db 341 VTS-----SFKH 347
 RESULT 7
 O9GZ04 PRELIMINARY; PRT; 415 AA.
 ID O9GZ04
 AC O9GZ04; O9NRAS;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Neuromedin U receptor 2 (Neuromedin U receptor-type 2) (G
 protein-coupled receptor TGR-1).
 DE NMUR2 OR NMUR2R OR TGR-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxId=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20490668; PubMed=10899166;
 RA Radatz R., Wilson A.E., Artyushyn R., Bonini J.A., Borowsky B.,
 RA Botelu L.W., Zhou S., Kouranova E.V., Nagorny R., Guevara M.S.,
 RA Dai M., Lerman G.S., Vayse P.J., Branche T.A., Gerald C., Forray C.,
 RA Adham N.;
 RT "Identification and Characterization of Two Neuromedin U Receptors
 RT Differentially Expressed in Peripheral Tissues and the Central Nervous
 RT System";
 RL J. Biol. Chem. 275:32452-32459 (2000).
 RN 12
 RP SEQUENCE FROM N.A.
 RX PubMed=11010960;
 RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
 RA Gustafson E.L., Monisma F.J., Jr., Hedrick J.A.,
 RT "Identification of a Novel Neuromedin U Receptor Subtype Expressed in
 RT the Central Nervous System";
 RL J. Biol. Chem. 275:39482-39486 (2000).
 RN 13
 RP SEQUENCE FROM N.A.
 RA Pang L., Wang S., Laz T., Hedrick J.A.;
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN 14
 RP SEQUENCE OF 4-415 FROM N.A.
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
 RA Stral J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 RT feeding";
 RL Nature 406:70-74 (2000).
 RN 15
 RP SEQUENCE OF 4-415 FROM N.A.
 RX PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shibata Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 RT Neuromedin U Receptor";
 RL J. Biol. Chem. 275:29528-29532 (2000).
 DR EMBL; AF272363; AAG24794.1; -
 DR EMBL; AF292402; AAG03064.1; -
 DR EMBL; AF242874; AAF82755.1; -
 DR EMBL; AB041228; BAB13721.1; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRKHODPSN.
 DR PRINTS; PR01565; NEUROMEDINUR.
 DR PRINTS; PR01567; NEUROMEDIN2R.
 DR PRINTS; PR01570; NPREFECREPOR.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.

DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW RECEPTOR.
 SQ SEQUENCE 415 AA; 44725 MW; C2BACD84B31390F CRC64;

Query Match 22.0%; Score 474; DB 4; Length 415;
 Best Local Similarity 30.8%; Pred. No. 7e-32;
 Matches 115; Conservative 78; Mismatches 98; Indels 82; Gaps 12;

QY 30 CSPFPLGALVPTAVCLCFVVGSGNVVTVMILGRYDMRTTNLYLGSMAVSDLLI-L 88
 DB 37 CGFRSRHFLPVSVAVALIFLVGSGNVLVCWIVHQTLPNTNYLFSLAUSDLLVL 96
 QY 89 IGLPDLRLMRSPRWVPGPLLCLSLYEGCCTVATLHMTALSYERYLAICRPLARV 148
 DB 97 LGMPLIEYEMHNYRPFLEPGVCYFETALFETVCFASTLSTVSERYVAIHPFRKL 156
 QY 149 LVTRRRVALLAVMAVALLSGPFLVGV--QDQGISVVPGLNGTARIASSPLASSP 206
 DB 157 QSTRRRALRLIGIVGSGVFLPSPNTSIHGIRKFFPNSLVPG-----SATCTYIK 208
 QY 207 PLMLSRAPPSPSPGPTAAALFSSRCRPPSPAQGLRVMLWTTAYEF-LPFLCLSI 265
 DB 209 PMWV-----YNLIQATSFLLPMTLISV 234
 QY 266 IYGLIGRLMSRRPLRPPASGRGRHQT-KRVLVVLVLAFTICWLP 313
 DB 235 LYVLMALRLKDKSLKDEADGNANIQRPCK-----KSNKMLFVLVFAICWAP 283
 QY 314 PHVGIYIYNTED-SRMWYFSQYFNIVAL--QLFYLSASINPLIYNLISKRYRAAFKL 369
 DB 284 FHIDRLFFSFEVSEWSESL--AAVENLVHVSGVFFYSSAVNPITNLSRRFQ-AAFQV 340
 QY 370 LLARKSRPGRGHR 382
 DB 341 VIS-----SFHK 347

RESULT 8

Q9JIB1 PRELIMINARY; PRT; 395 AA.

AC Q9JIB1; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Neuromedin U receptor 2.
 GN NMUR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Peng S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
 RA Strait J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McEee K.K.,
 RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austen C.P.,
 RA Caskey T., van der Ploeg I.H.T., Liu Q.,
 RT "Identification of receptors for neuromedin U and its role in
 feeding.";
 RL Nature 406:70-74 (2000).
 DR EMBL; AF242875; AAF82756.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PRO0237; GPCRHOOPS.
 DR PRINTS; PRO1565; NEUROMEDINUR.
 DR PRINTS; PRO1567; NEUROMEDINUR.
 DR PRINTS; PRO1570; NEUROMEDINUR.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 395 AA; 44722 MW; 01D3765B5D535C0 CRC64;

Query Match 21.9%; Score 472.5; DB 11; Length 395;
 Best Local Similarity 30.5%; Pred. No. 8.9e-32;
 Matches 111; Conservative 81; Mismatches 115; Indels 57; Gaps 9;

QY 30 CSPFPLGALVPTAVCLCFVVGSGNVVTVMILGRYDMRTTNLYLGSMAVSDLLI-L 88
 DB 29 CGFRSRHFLPVSVAVALIFLVGSGNVLVCWIVHQTLPNTNYLFSLAUSDLLVL 88
 QY 89 IGLPDLRLMRSPRWVPGPLLCLSLYEGCCTVATLHMTALSYERYLAICRPLARV 148
 DB 97 LGMPLIEYEMHNYRPFLEPGVCYFETALFETVCFASTLSTVSERYVAIHPFRKL 148
 QY 149 LVTRRRVALLAVMAVALLSGPFLVGV--QDQGISVVPGLNGTARIASSPLASSP 206
 DB 149 ESTRRARLRILSLVSVFSPNTSIHGIRKFFPNSLVPG-----SATCTYIK 200
 QY 207 PLMLSRAPPSPSPGPTAAALFSSRCRPPSPAQGLRVMLWTTAYEF-LPFLCLSI 265
 DB 201 PMWV-----YNLIQATSFLLPMTLISV 226
 QY 266 IYGLIGRLMSRRPLRPPASGRGRHQT-KRVLVVLVLAFTICWLP 324
 DB 227 LYVLMALRLKDKSLKDEADGNANIQRPCK-----KSNKMLFVLVFAICWAP 286
 QY 325 ED--SRMWYFSQYFNIVAL--QLFYLSASINPLIYNLISKRYRAAFKL 375
 DB 287 EHWTESLAAVNPLHVSGVFFYSSAVNPITNLSRRFQ-AAFQV 345
 QY 376 RPRG 379
 DB 346 RPOG 349

RESULT 9

Q9ESQ4 PRELIMINARY; PRT; 395 AA.

AC Q9ESQ4; 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE G-protein-coupled receptor 1GR-1.
 GN 1GR-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20449029; PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 Neuromedin U Receptor.";
 RL U. Biol. Chem. 275:29528-29532 (2000).
 DR EMBL; AB041229; BAB13722.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PRO0237; GPCRHOOPS.
 DR PRINTS; PRO1565; NEUROMEDINUR.
 DR PRINTS; PRO1567; NEUROMEDINUR.
 DR PRINTS; PRO1570; NEUROMEDINUR.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 395 AA; 44756 MW; EABD6D36B3C355DA CRC64;

Query Match 21.8%; Score 470.5; DB 11; Length 395;
 Best Local Similarity 31.2%; Pred. No. 1.3e-31;
 Matches 107; Conservative 75; Mismatches 112; Indels 49; Gaps 7;

QY 30 CSPFPLGALVPTAVCLCFVVGSGNVVTVMILGRYDMRTTNLYLGSMAVSDLLI-L 88

```

Db 29 CGPRSDLSLPSVAVALIFLVGMGNLLVCMVIVHQTLPNTNYFLSLAVSDLLVL 88
Qy 89 LGLPEDLRLMRSPWVFGPLICRLSLVYEGCTATLHMTALSVERYLAI CRPIRAV 148
Db 89 LGMPLEIEMHNYFLPGVCYFKTALFETVCFASISLVTVERVYALVHPRFRAK 148
Qy 149 LVTRRRVRLALVAVALLSAGPFLVGV--EODPGISVVPGLNGTARIASSPLASSP 206
Db 149 ESTRRRLRLISLWMSFVSFSLPNTSINGIKQHPNGSSVPG-----SATCTVK 200
Qy 207 PLMLSRAPPPSPSGPBTMAALFSGRCPSPAGLGRVMTLTTAYFF--LPFLCLSI 265
Db 201 PMWV-----YNLIQATSLFYLLPMTLISV 226
Qy 266 IYGLIGRELMSRRPLRGPASGRGRHROT-KRVLVAVLAFICMLPHVGRITIIYNT 324
Db 227 IYLMGLRLKRDESLANKVAVNIHPRSRKSVTKMLFVLVFAICWTPHVDRLFFSFV 286
Qy 325 ED-SRMWFSQYFNIVALQFLTSLASINPIYXLSKKYRAA 365
Db 287 EEMTESLAAVFNLIVHVSQVFFYLSAVNPFIYXLSRRFRRAA 329

RESULT 10
ID 055040 PRELIMINARY; PRT; 405 AA.
AC 055040;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Orphan G protein-coupled receptor.
GN GPR66.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA McKee K.K., Tan C.T., Liu J., Palyha O.C., Feigheimer S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D.,
RA Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF044602; AAC02681.1; -.
DR MGI; MGI:1341898; GPR66.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01566; NEUROMEDINUR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
KW SEQUENCE 405 AA; 45609 MW; F1BA493DBB81F34 CRC64;

Query Match 21.2%; Score 456.5; DB 11; Length 405;
Best Local Similarity 32.2%; Pred. No. 2e-30;
Matches 133; Conservative 62; Mismatches 141; Indels 77; Gaps 14;

```

```

Qy 250 WTTAYFFFLCLSLIYGLIGELMSRRPL-----RGPASG-----REGRHR 294
Db 211 TTALEFCLPMVITISVYLLIGLRLRRERMLQVEVGRKTAATQETSRRRIQDGR 270
Qy 295 QTRVULVYVLAITICMLPHVGRITIIY-----NTEDSRMWFSQYFNIVALQFLYS 349
Db 271 QVTKMLFALVVGICWAPHADRIIMSLVYGHSTBELHLAY--QCVHISGIFFYLSA 326
Qy 350 INPIYXLSKKYRAAFAKLL-----LARKSRPGRFHS-RDTAGVAGCTG 395
Db 329 ANPVLVSLMSTRFRFTFLQALGTCCHRRQPYHGSHNIRLTGTGTLCDVG 381

RESULT 11
ID 017239 PRELIMINARY; PRT; 418 AA.
AC 017239;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 47.7 kDa protein.
GN K10B4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=BRISTOL N2;
RA Sammons L., Wohlmann P., Antonion B.;
RT "The sequence of C. elegans cosmid K10B4."
RT submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF025463; AAB71009.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Hypothetical protein.
KW SEQUENCE 418 AA; 47735 MW; CE0416539CA3B27 CRC64;

Query Match 20.1%; Score 434; DB 5; Length 418;
Best Local Similarity 30.1%; Pred. No. 1.7e-28;
Matches 109; Conservative 77; Mismatches 138; Indels 38; Gaps 10;

```

QY 204 SSPLMLSRAPPSPGSPETAALFSSREC-----PPSPAQLGALRYMLVTTAYFEL 258
 Db 202 -----NRIFFPV-----TDGIFVLTHTFCAMNQSRPDQCK-----IIIFAVTFVFI 244
 QY 259 PFLCLSTLYGLIGRELMSSRRPLGPG--AAGRGREHGQTRKLVLVLAFLICMLPHV 316
 Db 245 PALAIVIMYAHIAVQLSSSEIDLGDKMKVKRRKNSKNTVLKMLLSVITFPCWLPFHI 304
 QY 317 GRITVINTEDSRMMYFSQYFNIVALQLFYL-----ASINPILYNLISKRYRAAPFL 369
 Db 305 QRLLSVYTTWSETTITSPVQFLMIVFYISGFCYSSAANPILYNLISQKRYSAFCRT 364
 QY 370 LL 371
 Db 365 IL 366

RESULT 12
 QYFN4 PRELIMINARY; PRT; 428 AA.
 AC QYFN4; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE CG9918 protein.
 GN CG9918.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foele C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kailush K.A., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazzolo M., Platten G.S., Pan S., Pollard D., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schebler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stepien M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao C., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).

CC -- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AB03703; AAF55016.1; --
 DR Flybase; FBgn0038201; CG9918.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PROSITE_1; 1.
 DR PROSITE; PROSITE_2; 1.
 DR PROSITE; PROSITE_3; 1.
 DR PROSITE; PROSITE_4; 1.
 DR PROSITE; PROSITE_5; 1.
 DR PROSITE; PROSITE_6; 1.
 DR PROSITE; PROSITE_7; 1.
 DR PROSITE; PROSITE_8; 1.
 DR PROSITE; PROSITE_9; 1.
 DR PROSITE; PROSITE_10; 1.
 DR PROSITE; PROSITE_11; 1.
 DR PROSITE; PROSITE_12; 1.
 DR PROSITE; PROSITE_13; 1.
 DR PROSITE; PROSITE_14; 1.
 DR PROSITE; PROSITE_15; 1.
 DR PROSITE; PROSITE_16; 1.
 DR PROSITE; PROSITE_17; 1.
 DR PROSITE; PROSITE_18; 1.
 DR PROSITE; PROSITE_19; 1.
 DR PROSITE; PROSITE_20; 1.
 DR PROSITE; PROSITE_21; 1.
 DR PROSITE; PROSITE_22; 1.
 DR PROSITE; PROSITE_23; 1.
 DR PROSITE; PROSITE_24; 1.
 DR PROSITE; PROSITE_25; 1.
 DR PROSITE; PROSITE_26; 1.
 DR PROSITE; PROSITE_27; 1.
 DR PROSITE; PROSITE_28; 1.
 DR PROSITE; PROSITE_29; 1.
 DR PROSITE; PROSITE_30; 1.
 DR PROSITE; PROSITE_31; 1.
 DR PROSITE; PROSITE_32; 1.
 DR PROSITE; PROSITE_33; 1.
 DR PROSITE; PROSITE_34; 1.
 DR PROSITE; PROSITE_35; 1.
 DR PROSITE; PROSITE_36; 1.
 DR PROSITE; PROSITE_37; 1.
 DR PROSITE; PROSITE_38; 1.
 DR PROSITE; PROSITE_39; 1.
 DR PROSITE; PROSITE_40; 1.
 DR PROSITE; PROSITE_41; 1.
 DR PROSITE; PROSITE_42; 1.
 DR PROSITE; PROSITE_43; 1.
 DR PROSITE; PROSITE_44; 1.
 DR PROSITE; PROSITE_45; 1.
 DR PROSITE; PROSITE_46; 1.
 DR PROSITE; PROSITE_47; 1.
 DR PROSITE; PROSITE_48; 1.
 DR PROSITE; PROSITE_49; 1.
 DR PROSITE; PROSITE_50; 1.
 DR PROSITE; PROSITE_51; 1.
 DR PROSITE; PROSITE_52; 1.
 DR PROSITE; PROSITE_53; 1.
 DR PROSITE; PROSITE_54; 1.
 DR PROSITE; PROSITE_55; 1.
 DR PROSITE; PROSITE_56; 1.
 DR PROSITE; PROSITE_57; 1.
 DR PROSITE; PROSITE_58; 1.
 DR PROSITE; PROSITE_59; 1.
 DR PROSITE; PROSITE_60; 1.
 DR PROSITE; PROSITE_61; 1.
 DR PROSITE; PROSITE_62; 1.
 DR PROSITE; PROSITE_63; 1.
 DR PROSITE; PROSITE_64; 1.
 DR PROSITE; PROSITE_65; 1.
 DR PROSITE; PROSITE_66; 1.
 DR PROSITE; PROSITE_67; 1.
 DR PROSITE; PROSITE_68; 1.
 DR PROSITE; PROSITE_69; 1.
 DR PROSITE; PROSITE_70; 1.
 DR PROSITE; PROSITE_71; 1.
 DR PROSITE; PROSITE_72; 1.
 DR PROSITE; PROSITE_73; 1.
 DR PROSITE; PROSITE_74; 1.
 DR PROSITE; PROSITE_75; 1.
 DR PROSITE; PROSITE_76; 1.
 DR PROSITE; PROSITE_77; 1.
 DR PROSITE; PROSITE_78; 1.
 DR PROSITE; PROSITE_79; 1.
 DR PROSITE; PROSITE_80; 1.
 DR PROSITE; PROSITE_81; 1.
 DR PROSITE; PROSITE_82; 1.
 DR PROSITE; PROSITE_83; 1.
 DR PROSITE; PROSITE_84; 1.
 DR PROSITE; PROSITE_85; 1.
 DR PROSITE; PROSITE_86; 1.
 DR PROSITE; PROSITE_87; 1.
 DR PROSITE; PROSITE_88; 1.
 DR PROSITE; PROSITE_89; 1.
 DR PROSITE; PROSITE_90; 1.
 DR PROSITE; PROSITE_91; 1.
 DR PROSITE; PROSITE_92; 1.
 DR PROSITE; PROSITE_93; 1.
 DR PROSITE; PROSITE_94; 1.
 DR PROSITE; PROSITE_95; 1.
 DR PROSITE; PROSITE_96; 1.
 DR PROSITE; PROSITE_97; 1.
 DR PROSITE; PROSITE_98; 1.
 DR PROSITE; PROSITE_99; 1.
 DR PROSITE; PROSITE_100; 1.
 DR PROSITE; PROSITE_101; 1.
 DR PROSITE; PROSITE_102; 1.
 DR PROSITE; PROSITE_103; 1.
 DR PROSITE; PROSITE_104; 1.
 DR PROSITE; PROSITE_105; 1.
 DR PROSITE; PROSITE_106; 1.
 DR PROSITE; PROSITE_107; 1.
 DR PROSITE; PROSITE_108; 1.
 DR PROSITE; PROSITE_109; 1.
 DR PROSITE; PROSITE_110; 1.
 DR PROSITE; PROSITE_111; 1.
 DR PROSITE; PROSITE_112; 1.
 DR PROSITE; PROSITE_113; 1.
 DR PROSITE; PROSITE_114; 1.
 DR PROSITE; PROSITE_115; 1.
 DR PROSITE; PROSITE_116; 1.
 DR PROSITE; PROSITE_117; 1.
 DR PROSITE; PROSITE_118; 1.
 DR PROSITE; PROSITE_119; 1.
 DR PROSITE; PROSITE_120; 1.
 DR PROSITE; PROSITE_121; 1.
 DR PROSITE; PROSITE_122; 1.
 DR PROSITE; PROSITE_123; 1.
 DR PROSITE; PROSITE_124; 1.
 DR PROSITE; PROSITE_125; 1.
 DR PROSITE; PROSITE_126; 1.
 DR PROSITE; PROSITE_127; 1.
 DR PROSITE; PROSITE_128; 1.
 DR PROSITE; PROSITE_129; 1.
 DR PROSITE; PROSITE_130; 1.
 DR PROSITE; PROSITE_131; 1.
 DR PROSITE; PROSITE_132; 1.
 DR PROSITE; PROSITE_133; 1.
 DR PROSITE; PROSITE_134; 1.
 DR PROSITE; PROSITE_135; 1.
 DR PROSITE; PROSITE_136; 1.
 DR PROSITE; PROSITE_137; 1.
 DR PROSITE; PROSITE_138; 1.
 DR PROSITE; PROSITE_139; 1.
 DR PROSITE; PROSITE_140; 1.
 DR PROSITE; PROSITE_141; 1.
 DR PROSITE; PROSITE_142; 1.
 DR PROSITE; PROSITE_143; 1.
 DR PROSITE; PROSITE_144; 1.
 DR PROSITE; PROSITE_145; 1.
 DR PROSITE; PROSITE_146; 1.
 DR PROSITE; PROSITE_147; 1.
 DR PROSITE; PROSITE_148; 1.
 DR PROSITE; PROSITE_149; 1.
 DR PROSITE; PROSITE_150; 1.
 DR PROSITE; PROSITE_151; 1.
 DR PROSITE; PROSITE_152; 1.
 DR PROSITE; PROSITE_153; 1.
 DR PROSITE; PROSITE_154; 1.
 DR PROSITE; PROSITE_155; 1.
 DR PROSITE; PROSITE_156; 1.
 DR PROSITE; PROSITE_157; 1.
 DR PROSITE; PROSITE_158; 1.
 DR PROSITE; PROSITE_159; 1.
 DR PROSITE; PROSITE_160; 1.
 DR PROSITE; PROSITE_161; 1.
 DR PROSITE; PROSITE_162; 1.
 DR PROSITE; PROSITE_163; 1.
 DR PROSITE; PROSITE_164; 1.
 DR PROSITE; PROSITE_165; 1.
 DR PROSITE; PROSITE_166; 1.
 DR PROSITE; PROSITE_167; 1.
 DR PROSITE; PROSITE_168; 1.
 DR PROSITE; PROSITE_169; 1.
 DR PROSITE; PROSITE_170; 1.
 DR PROSITE; PROSITE_171; 1.
 DR PROSITE; PROSITE_172; 1.
 DR PROSITE; PROSITE_173; 1.
 DR PROSITE; PROSITE_174; 1.
 DR PROSITE; PROSITE_175; 1.
 DR PROSITE; PROSITE_176; 1.
 DR PROSITE; PROSITE_177; 1.
 DR PROSITE; PROSITE_178; 1.
 DR PROSITE; PROSITE_179; 1.
 DR PROSITE; PROSITE_180; 1.
 DR PROSITE; PROSITE_181; 1.
 DR PROSITE; PROSITE_182; 1.
 DR PROSITE; PROSITE_183; 1.
 DR PROSITE; PROSITE_184; 1.
 DR PROSITE; PROSITE_185; 1.
 DR PROSITE; PROSITE_186; 1.
 DR PROSITE; PROSITE_187; 1.
 DR PROSITE; PROSITE_188; 1.
 DR PROSITE; PROSITE_189; 1.
 DR PROSITE; PROSITE_190; 1.
 DR PROSITE; PROSITE_191; 1.
 DR PROSITE; PROSITE_192; 1.
 DR PROSITE; PROSITE_193; 1.
 DR PROSITE; PROSITE_194; 1.
 DR PROSITE; PROSITE_195; 1.
 DR PROSITE; PROSITE_196; 1.
 DR PROSITE; PROSITE_197; 1.
 DR PROSITE; PROSITE_198; 1.
 DR PROSITE; PROSITE_199; 1.
 DR PROSITE; PROSITE_200; 1.
 DR PROSITE; PROSITE_201; 1.
 DR PROSITE; PROSITE_202; 1.
 DR PROSITE; PROSITE_203; 1.
 DR PROSITE; PROSITE_204; 1.
 DR PROSITE; PROSITE_205; 1.
 DR PROSITE; PROSITE_206; 1.
 DR PROSITE; PROSITE_207; 1.
 DR PROSITE; PROSITE_208; 1.
 DR PROSITE; PROSITE_209; 1.
 DR PROSITE; PROSITE_210; 1.
 DR PROSITE; PROSITE_211; 1.
 DR PROSITE; PROSITE_212; 1.
 DR PROSITE; PROSITE_213; 1.
 DR PROSITE; PROSITE_214; 1.
 DR PROSITE; PROSITE_215; 1.
 DR PROSITE; PROSITE_216; 1.
 DR PROSITE; PROSITE_217; 1.
 DR PROSITE; PROSITE_218; 1.
 DR PROSITE; PROSITE_219; 1.
 DR PROSITE; PROSITE_220; 1.
 DR PROSITE; PROSITE_221; 1.
 DR PROSITE; PROSITE_222; 1.
 DR PROSITE; PROSITE_223; 1.
 DR PROSITE; PROSITE_224; 1.
 DR PROSITE; PROSITE_225; 1.
 DR PROSITE; PROSITE_226; 1.
 DR PROSITE; PROSITE_227; 1.
 DR PROSITE; PROSITE_228; 1.
 DR PROSITE; PROSITE_229; 1.
 DR PROSITE; PROSITE_230; 1.
 DR PROSITE; PROSITE_231; 1.
 DR PROSITE; PROSITE_232; 1.
 DR PROSITE; PROSITE_233; 1.
 DR PROSITE; PROSITE_234; 1.
 DR PROSITE; PROSITE_235; 1.
 DR PROSITE; PROSITE_236; 1.
 DR PROSITE; PROSITE_237; 1.
 DR PROSITE; PROSITE_238; 1.
 DR PROSITE; PROSITE_239; 1.
 DR PROSITE; PROSITE_240; 1.
 DR PROSITE; PROSITE_241; 1.
 DR PROSITE; PROSITE_242; 1.
 DR PROSITE; PROSITE_243; 1.
 DR PROSITE; PROSITE_244; 1.
 DR PROSITE; PROSITE_245; 1.
 DR PROSITE; PROSITE_246; 1.
 DR PROSITE; PROSITE_247; 1.
 DR PROSITE; PROSITE_248; 1.
 DR PROSITE; PROSITE_249; 1.
 DR PROSITE; PROSITE_250; 1.
 DR PROSITE; PROSITE_251; 1.
 DR PROSITE; PROSITE_252; 1.
 DR PROSITE; PROSITE_253; 1.
 DR PROSITE; PROSITE_254; 1.
 DR PROSITE; PROSITE_255; 1.
 DR PROSITE; PROSITE_256; 1.
 DR PROSITE; PROSITE_257; 1.
 DR PROSITE; PROSITE_258; 1.
 DR PROSITE; PROSITE_259; 1.
 DR PROSITE; PROSITE_260; 1.
 DR PROSITE; PROSITE_261; 1.
 DR PROSITE; PROSITE_262; 1.
 DR PROSITE; PROSITE_263; 1.
 DR PROSITE; PROSITE_264; 1.
 DR PROSITE; PROSITE_265; 1.
 DR PROSITE; PROSITE_266; 1.
 DR PROSITE; PROSITE_267; 1.
 DR PROSITE; PROSITE_268; 1.
 DR PROSITE; PROSITE_269; 1.
 DR PROSITE; PROSITE_270; 1.
 DR PROSITE; PROSITE_271; 1.
 DR PROSITE; PROSITE_272; 1.
 DR PROSITE; PROSITE_273; 1.
 DR PROSITE; PROSITE_274; 1.
 DR PROSITE; PROSITE_275; 1.
 DR PROSITE; PROSITE_276; 1.
 DR PROSITE; PROSITE_277; 1.
 DR PROSITE; PROSITE_278; 1.
 DR PROSITE; PROSITE_279; 1.
 DR PROSITE; PROSITE_280; 1.
 DR PROSITE; PROSITE_281; 1.
 DR PROSITE; PROSITE_282; 1.
 DR PROSITE; PROSITE_283; 1.
 DR PROSITE; PROSITE_284; 1.
 DR PROSITE; PROSITE_285; 1.
 DR PROSITE; PROSITE_286; 1.
 DR PROSITE; PROSITE_287; 1.
 DR PROSITE; PROSITE_288; 1.
 DR PROSITE; PROSITE_289; 1.
 DR PROSITE; PROSITE_290; 1.
 DR PROSITE; PROSITE_291; 1.
 DR PROSITE; PROSITE_292; 1.
 DR PROSITE; PROSITE_293; 1.
 DR PROSITE; PROSITE_294; 1.
 DR PROSITE; PROSITE_295; 1.
 DR PROSITE; PROSITE_296; 1.
 DR PROSITE; PROSITE_297; 1.
 DR PROSITE; PROSITE_298; 1.
 DR PROSITE; PROSITE_299; 1.
 DR PROSITE; PROSITE_300; 1.
 DR PROSITE; PROSITE_301; 1.
 DR PROSITE; PROSITE_302; 1.
 DR PROSITE; PROSITE_303; 1.
 DR PROSITE; PROSITE_304; 1.
 DR PROSITE; PROSITE_305; 1.
 DR PROSITE; PROSITE_306; 1.
 DR PROSITE; PROSITE_307; 1.
 DR PROSITE; PROSITE_308; 1.
 DR PROSITE; PROSITE_309; 1.
 DR PROSITE; PROSITE_310; 1.
 DR PROSITE; PROSITE_311; 1.
 DR PROSITE; PROSITE_312; 1.
 DR PROSITE; PROSITE_313; 1.
 DR PROSITE; PROSITE_314; 1.
 DR PROSITE; PROSITE_315; 1.
 DR PROSITE; PROSITE_316; 1.
 DR PROSITE; PROSITE_317; 1.
 DR PROSITE; PROSITE_318; 1.
 DR PROSITE; PROSITE_319; 1.
 DR PROSITE; PROSITE_320; 1.
 DR PROSITE; PROSITE_321; 1.
 DR PROSITE; PROSITE_322; 1.
 DR PROSITE; PROSITE_323; 1.
 DR PROSITE; PROSITE_324; 1.
 DR PROSITE; PROSITE_325; 1.
 DR PROSITE; PROSITE_326; 1.
 DR PROSITE; PROSITE_327; 1.
 DR PROSITE; PROSITE_328; 1.
 DR PROSITE; PROSITE_329; 1.
 DR PROSITE; PROSITE_330; 1.
 DR PROSITE; PROSITE_331; 1.
 DR PROSITE; PROSITE_332; 1.
 DR PROSITE; PROSITE_333; 1.
 DR PROSITE; PROSITE_334; 1.
 DR PROSITE; PROSITE_335; 1.
 DR PROSITE; PROSITE_336; 1.
 DR PROSITE; PROSITE_337; 1.
 DR PROSITE; PROSITE_338; 1.
 DR PROSITE; PROSITE_339; 1.
 DR PROSITE; PROSITE_340; 1.
 DR PROSITE; PROSITE_341; 1.
 DR PROSITE; PROSITE_342; 1.
 DR PROSITE; PROSITE_343; 1.
 DR PROSITE; PROSITE_344; 1.
 DR PROSITE; PROSITE_345; 1.
 DR PROSITE; PROSITE_346; 1.
 DR PROSITE; PROSITE_347; 1.
 DR PROSITE; PROSITE_348; 1.
 DR PROSITE; PROSITE_349; 1.
 DR PROSITE; PROSITE_350; 1.
 DR PROSITE; PROSITE_351; 1.
 DR PROSITE; PROSITE_352; 1.
 DR PROSITE; PROSITE_353; 1.
 DR PROSITE; PROSITE_354; 1.
 DR PROSITE; PROSITE_355; 1.
 DR PROSITE; PROSITE_356; 1.
 DR PROSITE; PROSITE_357; 1.
 DR PROSITE; PROSITE_358; 1.
 DR PROSITE; PROSITE_359; 1.
 DR PROSITE; PROSITE_360; 1.
 DR PROSITE; PROSITE_361; 1.
 DR PROSITE; PROSITE_362; 1.
 DR PROSITE; PROSITE_363; 1.
 DR PROSITE; PROSITE_364; 1.
 DR PROSITE; PROSITE_365; 1.
 DR PROSITE; PROSITE_366; 1.
 DR PROSITE; PROSITE_367; 1.
 DR PROSITE; PROSITE_368; 1.
 DR PROSITE; PROSITE_369; 1.
 DR PROSITE; PROSITE_370; 1.
 DR PROSITE; PROSITE_371; 1.
 DR PROSITE; PROSITE_372; 1.
 DR PROSITE; PROSITE_373; 1.
 DR PROSITE; PROSITE_374; 1.
 DR PROSITE; PROSITE_375; 1.
 DR PROSITE; PROSITE_376; 1.
 DR PROSITE; PROSITE_377; 1.
 DR PROSITE; PROSITE_378; 1.
 DR PROSITE; PROSITE_379; 1.
 DR PROSITE; PROSITE_380; 1.
 DR PROSITE; PROSITE_381; 1.
 DR PROSITE; PROSITE_382; 1.
 DR PROSITE; PROSITE_383; 1.
 DR PROSITE; PROSITE_384; 1.
 DR PROSITE; PROSITE_385; 1.
 DR PROSITE; PROSITE_386; 1.
 DR PROSITE; PROSITE_387; 1.
 DR PROSITE; PROSITE_388; 1.
 DR PROSITE; PROSITE_389; 1.
 DR PROSITE; PROSITE_390; 1.
 DR PROSITE; PROSITE_391; 1.
 DR PROSITE; PROSITE_392; 1.
 DR PROSITE; PROSITE_393; 1.
 DR PROSITE; PROSITE_394; 1.
 DR PROSITE; PROSITE_395; 1.
 DR PROSITE; PROSITE_396; 1.
 DR PROSITE; PROSITE_397; 1.
 DR PROSITE; PROSITE_398; 1.
 DR PROSITE; PROSITE_399; 1.
 DR PROSITE; PROSITE_400; 1.
 DR PROSITE; PROSITE_401; 1.
 DR PROSITE; PROSITE_402; 1.
 DR PROSITE; PROSITE_403; 1.
 DR PROSITE; PROSITE_404; 1.
 DR PROSITE; PROSITE_405; 1.
 DR PROSITE; PROSITE_406; 1.
 DR PROSITE; PROSITE_407; 1.
 DR PROSITE; PROSITE_408; 1.
 DR PROSITE; PROSITE_409; 1.
 DR PROSITE; PROSITE_410; 1.
 DR PROSITE; PROSITE_411; 1.
 DR PROSITE; PROSITE_412; 1.
 DR PROSITE; PROSITE_413; 1.
 DR PROSITE; PROSITE_414; 1.
 DR PROSITE; PROSITE_415; 1.
 DR PROSITE; PROSITE_416; 1.
 DR PROSITE; PROSITE_417; 1.
 DR PROSITE; PROSITE_418; 1.
 DR PROSITE; PROSITE_419; 1.
 DR PROSITE; PROSITE_420; 1.
 DR PROSITE; PROSITE_421; 1.
 DR PROSITE; PROSITE_422; 1.
 DR PROSITE; PROSITE_423; 1.
 DR PROSITE; PROSITE_424; 1.
 DR PROSITE; PROSITE_425; 1.
 DR PROSITE; PROSITE_426; 1.
 DR PROSITE; PROSITE_427; 1.
 DR PROSITE; PROSITE_428; 1.
 DR PROSITE; PROSITE_429; 1.
 DR PROSITE; PROSITE_430; 1.
 DR PROSITE; PROSITE_431; 1.
 DR PROSITE; PROSITE_432; 1.
 DR PROSITE; PROSITE_433; 1.
 DR PROSITE; PROSITE_434; 1.
 DR PROSITE; PROSITE_435; 1.
 DR PROSITE; PROSITE_436; 1.
 DR PROSITE; PROSITE_437; 1.
 DR PROSITE; PROSITE_438; 1.
 DR PROSITE; PROSITE_439; 1.
 DR PROSITE; PROSITE_440; 1.
 DR PROSITE; PROSITE_441; 1.
 DR PROSITE; PROSITE_442; 1.
 DR PROSITE; PROSITE_443; 1.
 DR PROSITE; PROSITE_444; 1.
 DR PROSITE; PROSITE_445; 1.
 DR PROSITE; PROSITE_446; 1.
 DR PROSITE; PROSITE_447; 1.
 DR PROSITE; PROSITE_448; 1.
 DR PROSITE; PROSITE_449; 1.
 DR PROSITE; PROSITE_450; 1.
 DR PROSITE; PROSITE_451; 1.
 DR PROSITE; PROSITE_452; 1.
 DR PROSITE; PROSITE_453; 1.
 DR PROSITE; PROSITE_454; 1.
 DR PROSITE; PROSITE_455; 1.
 DR PROSITE; PROSITE_456; 1.
 DR PROSITE; PROSITE_457; 1.
 DR PROSITE; PROSITE_458; 1.
 DR PROSITE; PROSITE_459; 1.
 DR PROSITE; PROSITE_460; 1.
 DR PROSITE; PROSITE_461; 1.
 DR PROSITE; PROSITE_462; 1.
 DR PROSITE; PROSITE_463; 1.
 DR PROSITE; PROSITE_464; 1.
 DR PROSITE; PROSITE_465; 1.
 DR PROSITE; PROSITE_466; 1.
 DR PROSITE; PROSITE_467; 1.
 DR PROSITE; PROSITE_468; 1.
 DR PROSITE; PROSITE_469; 1.
 DR PROSITE; PROSITE_470; 1.
 DR PROSITE; PROSITE_471; 1.
 DR PROSITE; PROSITE_472; 1.
 DR PROSITE; PROSITE_473; 1.
 DR PROSITE; PROSITE_474; 1.
 DR PROSITE; PROSITE_475; 1.
 DR PROSITE; PROSITE_476; 1.
 DR PROSITE; PROSITE_477; 1.
 DR PROSITE; PROSITE_478; 1.
 DR PROSITE; PROSITE_479; 1.
 DR PROSITE; PROSITE_480; 1.
 DR PROSITE; PROSITE_481; 1.
 DR PROSITE; PROSITE_482; 1.
 DR PROSITE; PROSITE_483; 1.
 DR PROSITE; PROSITE_484; 1.
 DR PROSITE; PROSITE_485; 1.
 DR PROSITE; PROSITE_486; 1.
 DR PROSITE; PROSITE_487; 1.
 DR PROSITE; PROSITE_488; 1.
 DR PROSITE; PROSITE_489; 1.
 DR PROSITE; PROSITE_490; 1.
 DR PROSITE; PROSITE_491; 1.
 DR PROSITE; PROSITE_492; 1.
 DR PROSITE; PROSITE_493; 1.
 DR PROSITE; PROSITE_494; 1.
 DR PROSITE; PROSITE_495; 1.
 DR PROSITE; PROSITE_496; 1.
 DR PROSITE; PROSITE_497; 1.
 DR PROSITE; PROSITE_498; 1.
 DR PROSITE; PROSITE_499; 1.
 DR PROSITE; PROSITE_500; 1.
 DR PROSITE; PROSITE_501; 1.
 DR PROSITE; PROSITE_502; 1.
 DR PROSITE; PROSITE_503; 1.
 DR PROSITE; PROSITE_504; 1.
 DR PROSITE; PROSITE_505; 1.
 DR PROSITE; PROSITE_506; 1.
 DR PROSITE; PROSITE_507; 1.
 DR PROSITE; PROSITE_508; 1.
 DR PROSITE; PROSITE_509; 1.
 DR PROSITE; PROSITE_510; 1.
 DR PROSITE; PROSITE_511; 1.
 DR PROSITE; PROSITE_512; 1.
 DR PROSITE; PROSITE_513; 1.
 DR PROSITE; PROSITE_514; 1.
 DR PROSITE; PROSITE_515; 1.
 DR PROSITE; PROSITE_516; 1.
 DR PROSITE; PROSITE_517; 1.
 DR PROSITE; PROSITE_518; 1.
 DR PROSITE; PROSITE_519; 1.
 DR PROSITE; PROSITE_520; 1.
 DR PROSITE; PROSITE_521; 1.
 DR PROSITE; PROSITE_522; 1.
 DR PROSITE; PROSITE_523; 1.
 DR PROSITE; PROSITE_524; 1.
 DR PROSITE; PROSITE_525; 1.
 DR PROSITE; PROSITE_526; 1.
 DR PROSITE; PROSITE_527; 1.
 DR PROSITE; PROSITE_528; 1.
 DR PROSITE; PROSITE_529; 1.
 DR PROSITE; PROSITE_530; 1.
 DR PROSITE; PROSITE_531; 1.
 DR PROSITE; PROSITE_532; 1.
 DR PROSITE; PROSITE_533; 1.
 DR PROSITE; PROSITE_534; 1.
 DR PROSITE; PROSITE_535; 1.
 DR PROSITE; PROSITE_536; 1.
 DR PROSITE; PROSITE_537; 1.
 DR PROSITE; PROSITE_538; 1.
 DR PROSITE; PROSITE_539; 1.
 DR PROSITE; PROSITE_540; 1.
 DR PROSITE; PROSITE_541; 1.
 DR PROSITE; PROSITE_542; 1.
 DR PROSITE; PROSITE_543; 1.
 DR PROSITE; PROSITE_544; 1.
 DR PROSITE; PROSITE_545; 1.
 DR PROSITE; PROSITE_546; 1.
 DR PROSITE; PROSITE_547; 1.
 DR PROSITE; PROSITE_548; 1.
 DR PROSITE; PROSITE_549; 1.
 DR PROSITE; PROSITE_550; 1.
 DR PROSITE; PROSITE_551; 1.
 DR PROSITE; PROSITE_552; 1.
 DR PROSITE; PROSITE_553; 1.
 DR PROSITE; PROSITE_554; 1.
 DR PROSITE; PROSITE_555; 1.
 DR PROSITE; PROSITE_556; 1.
 DR PROSITE; PROSITE_557; 1.
 DR PROSITE; PROSITE_558; 1.
 DR PROSITE; PROSITE_559; 1.
 DR PROSITE; PROSITE_560; 1.
 DR PROSITE; PROSITE_561; 1.
 DR PROSITE; PROSITE_562; 1.
 DR PROSITE; PROSITE_563; 1.
 DR PROSITE; PROSITE_564; 1.
 DR PROSITE; PROSITE_565; 1.
 DR PROSITE; PROSITE_566; 1.
 DR PROSITE; PROSITE_567; 1.
 DR PROSITE; PROSITE_568; 1.
 DR PROSITE; PROSITE_569; 1.
 DR PROSITE; PROSITE_570; 1.
 DR PROSITE; PROSITE_571; 1.
 DR PROSITE; PROSITE_572; 1.
 DR PROSITE; PROSITE_573; 1.
 DR PROSITE; PROSITE_574; 1.
 DR PROSITE; PROSITE_575; 1.
 DR PROSITE; PROSITE_576; 1.
 DR PROSITE; PROSITE_577; 1.
 DR PROSITE; PROSITE_578; 1.
 DR PROSITE; PROSITE_579; 1.
 DR PROSITE; PROSITE_580

Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abrell J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Bailew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies S.,
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Folsler C., Gabrielian A.E., Garg N.S., Guan P., Harris M.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jajala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., Moshrefi A.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Munz D.W., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.V., Reese M.G.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*,"
 Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 RT -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AF003699; AAF54929.1; -
 DR Flybase: FBgn0038139; CG8795.
 DR InterPro: IPR000923; Bluecut_1.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHOPOPSN.
 DR PRINTS: PR01565; NEUROMEDINUR.
 DR PRINTS: PR00196; COPPER BLUE.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR G-protein coupled receptor; Glycoprotein; Transmembrane.
 KW SEQUENCE 595 AA; 65091 MW; 3E8E183BADF5471E CRC64;
 SO

Query Match 19.2%; Score 414; DB 5; Length 595;
 Best Local Similarity 29.7%; Pred. No. 1.2e-26;
 Matches 121; Conservative 61; Mismatches 138; Indels 88; Gaps 13;

DB 301 AYVSGTSGIESQWENFVSLIDYTSGLVFLSTCINPLDLYNIMSHKR-EAFVTLARHF 359
 QY 374 ---KRRPGFPHRS-----RDTGGEVAGDGTGCTGCTGTESSANKVWG 412
 DB 360 GLGCKKQGRGLPHTYSGALRRNQGSLRLH-TDSVRTTMTSMATTGG 406

RESULT 14
 ID Q9JIB2 PRELIMINARY; RT; 402 AA.
 AC Q9JIB2;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Neuremedin u receptor 1.
 GN NMUR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20351041; PubMed=10994543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 Zeng Z., Williams D.L., Reigner S.D., Nunes C.N., Murphy B.,
 Strick J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 Hrensky D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
 Chaskey T., van der Ploeg L.H.T., Liu Q.;
 RA "Identification of receptors for neuremedin U and its role in
 feeding,"
 RT
 RT Nature 406:70-74 (2000).
 RL Nature 406:70-74 (2000).
 DR EMBL: AF242873; AAF82754.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHOPOPSN.
 DR PRINTS: PR01565; NEUROMEDINUR.
 DR PRINTS: PR01566; NEUROMEDINUR.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
 KW Receptor.
 KW SEQUENCE 402 AA; 45782 MW; CB576083D9F7A095 CRC64;
 SO

Query Match 19.0%; Score 410; DB 11; Length 402;
 Best Local Similarity 29.9%; Pred. No. 1.7e-26;
 Matches 112; Conservative 68; Mismatches 130; Indels 64; Gaps 9;

Db 348 LGJGRCCHROPR 361

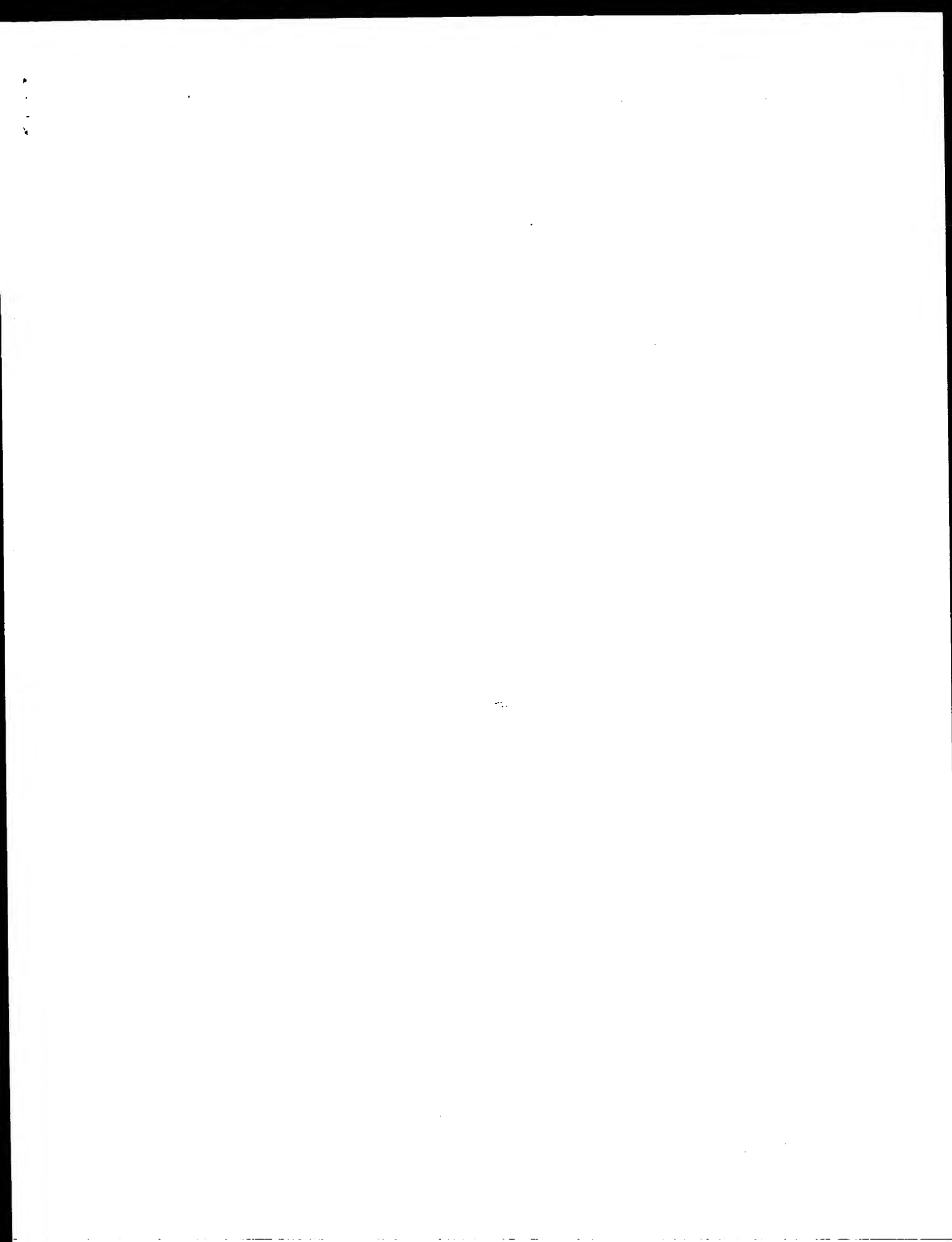
RESULT 15

08TBH6 PRELIMINARY; PRT; 410 AA.
 AC 08TBH6;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 OS Neurotensin receptor 2.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC022501; AAH22501.1; -
 KW Receptor.
 SQ SEQUENCE 410 AA; 45459 MW; DC3ADA4B6DS8D7BD CRC64;

Query Match 18.8%; Score 405; DB 4; Length 410;
 Best Local Similarity 30.9%; Pred. No. 4.5e-26;
 Matches 128; Conservative 61; Mismatches 135; Indels 90; Gaps 16;

QY 17 PMPALP---CDERCSPPPLGALVPTVCLTFVVGSGNVVTMLI-----GR 65
 DB 8 PPRSSNPGLSDARLGVDTRIMAKVFTALYALIMLGAAGNALSVHVLKARAGRGR 67
 QY 66 YRDMRTTNLYGMAVSDLLIL-GLPDLXR-LMRSRPWFEGPLLCRLSLVYEGCTY 123
 DB 68 LRH-----HYLSLALAGILLLVGPVELYSFVWFHYPMVFGDLGCRGYFVHELCA 120
 QY 124 ATLLHMTALSVRYLAICRPLRARVTRRRVALLVMAVALISGPFLLVGV----- 179
 DB 121 ATVLVAGLSAERCLAVCOPLRAKSLTPRTTRMLVALSWAASLGLMPMAVIMGQKHEL 180
 QY 180 -----EQDGIISVPGINGTARIASS-----PLASSPPL-----WLS 211
 DB 181 ETADGEPBPASRYCTVL--VSRVALQVFIQVNVLVSVPLALTAFLNGVTSHLALCS 238
 QY 212 RAPPPSPGSPETAAALPFRRCRPSPAOLGALRVMLVTTAYFFLPFLCLSLYGLIG 271
 DB 239 QVPSTSTP-GSSTPSRLLEISE-----GLISFIWKKT-----FIG--G 276
 QY 272 RELWSSRRPLRGPAASGRGRGHROTREVLLVVLAFICWLPBVGRITINTEDSR--- 328
 DB 277 QVSLVRHKDVR-----RIRSLQRSQVLRATVVMYVICMPLPHARLMYCYVDDAMTD 330
 QY 329 -MMYFSQYFNIALQLFYISASINPLILYNLSKKYRAAAFKLLARKSRPRGFH 381
 DB 331 PLYNFYHYFVMTNTLFYVSSAVTPLLNVAVSSFR-----RLFLNAVSSLCGEH 380

Search completed: March 16, 2003, 15:05:03
 Job time : 37 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:10 ; Search time 13 Seconds
(without alignments)
1314.481 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156
Sequence: 1 MGSPMNGSDGPEGAREPWP.....DTGGDTVGTETSANVKTMG 412

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2149	99.7	412	1	MTLR_HUMAN
2	858.5	39.8	366	1	GHSR_PIG
3	857.5	39.8	366	1	GHSR_HUMAN
4	855.5	39.7	364	1	GHSR_PAT
5	614	28.5	257	1	GHSR_MOUSE
6	498	23.1	424	1	NTR1_PAT
7	495.5	23.0	424	1	NTR1_MOUSE
8	491	22.8	418	1	NTR1_HUMAN
9	407	18.9	410	1	NTR2_HUMAN
10	404	18.7	417	1	NTR2_MOUSE
11	401	18.6	416	1	NTR2_PAT
12	369.5	17.1	362	1	SSRS_MOUSE
13	358	16.6	363	1	SSRS_PAT
14	356.5	16.5	418	1	SSRS_HUMAN
15	349.5	16.2	453	1	GP39_HUMAN
16	343	15.9	364	1	SSRS_HUMAN
17	342.5	15.9	380	1	OPRX_CAVPO
18	341.5	15.8	370	1	OPRX_HUMAN
19	338	15.7	380	1	OPRX_PAT
20	334	15.5	380	1	OPRX_HUMAN
21	334	15.5	388	1	SSR4_HUMAN
22	333.5	15.5	367	1	OPRX_MOUSE
23	333.5	15.5	367	1	OPRX_PAT
24	333.5	15.5	370	1	OPRX_PAT
25	333.5	15.5	428	1	SSR3_PAT
26	332.5	15.4	372	1	OPRX_HUMAN
27	331.5	15.4	395	1	TRFR_CHICK
28	331	15.4	428	1	SSR3_PAT
29	330.5	15.3	428	1	SSR3_MOUSE
30	329.5	15.3	370	1	OPRX_HUMAN
31	328	15.2	380	1	OPRX_MOUSE
32	326.5	15.1	371	1	GALS_MOUSE
33	326.5	15.1	384	1	SSR4_PAT

ALIGNMENTS

RESULT 1	ID	MTLR_HUMAN	STANDARD	PRT	412 AA.
AC	043153				
DT	15-DEC-1998	(Rel. 37, Created)			
DT	15-DEC-1998	(Rel. 37, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Motilin receptor (G-protein-coupled receptor GPR38).				
GN	MTLR1 OR MTLR OR GPR38.				
OS	Homo sapiens (Human).				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RX	MEDLINE=98110578; PubMed=9441746;				
RA	McGee K.K., Tan C.P., Palyna O.C., Liu J., Feighner S.D.,				
RA	Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;				
RT	"Cloning and characterization of two human G protein-coupled receptor				
RT	genes (GPR38 and GPR39) related to the growth hormone secretagogue				
RT	and neurotensin receptors.";				
RL	Genomics 46:426-434(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORMS A AND B).				
RX	MEDLINE=99316084; PubMed=10381885;				
RA	Feighner S.D., Tan C.P., McGee K.K., Palyna O.C., Hreniuk D.L.,				
RA	Pong S.-S., Austin C.P., Figueroa D., MacNeil D., Cascieri M.A.,				
RA	Nargund R., Bakhshi R., Abramowitz M., Soccio R., Kargman S.,				
RA	O'Neill G., van der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,				
RT	Howard A.D.,				
RT	"Receptor for motilin identified in the human gastrointestinal				
RT	system.";				
RL	Science 284:2184-2188(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM A).				
RA	Wall M.				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	FUNCTION.				
RX	MEDLINE=21219832; PubMed=11322507;				
RA	Smith R.G., Leonard R., Bailey A.R.T., Palyna O.C., Feighner S.D.,				
RA	Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;				
RT	"Growth hormone secretagogue receptor family members and ligands.";				
RL	Endocrine 14:9-14(2001).				
CC	-1- FUNCTION: Receptor for motilin.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are				
CC	produced by alternative splicing.				
CC	-1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE				
CC	MARROW.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				

FT DISULFID 116 198 BY SIMILARITY.
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 266 289 AVVFAFLLCMLPFHVRGYFSKS -> GGSQCLLESLPG
 FT VARSPLIC 266 289 PLHSCCLFSSP (IN ISOFORM 1B).
 FT VARSPLIC 290 366 MISSING (IN ISOFORM 1B).
 SQ SEQUENCE 366 AA, 41194 MW, 2C850B35F61B7C1C CRC64;
 Query Match 39.8%; Score 858.5; DB 1; Length 366;
 Best Local Similarity 44.9%; Pred. No. 1.6e-48;
 Matches 188; Conservative 55; Mismatches 105; Indels 71; Gaps 10;
 QY 5 NMGSDGDEGA-----REPMPALPCCD---ERRCSPPGLALVVTAVVCLCFVGVGSGN 56
 DB 2 WATPEEPEEPNLTLPDLGMDAPPENDSLVEELLPLPPTPLGAVTATCAVAFVAGIAGN 61
 QY 57 VVTVMILGCRDRTTNTLYLGSMAVSDLLTLGLPFDLYRMRSPWVGILLCRSLV 116
 DB 62 LITMLVVSFRERTTNTLYLSMAVSDLLTLGLPFDLYRMRSPWVGILLCRSLV 121
 QY 117 VEEGCTATLMTALSVERYLAICRPLRVLVTRRRVAVLAIVAAVALLSAGPFL 176
 DB 122 VSESCYATVLTALSVERYFAICRPLRVLVTRRRVAVLAIVAAVALLSAGPFL 181
 QY 177 VGEVDPGISVPLGNGTARINASSPLASSPMLSRAPPSPPSGPTAAALFSECR 236
 DB 182 VGEVD-----NGT-----DPRD-----TNECR 199
 QY 237 PS-PAOLGALRMLVTTAVFPLCLSLYLGLIGRELMSRRPLRGAASG---RER 291
 DB 200 ATEFAVRSGLLVVMVSSVFVFLPFCULTVLSLGRKM---RRKRGAAVSSLRDQ 256
 QY 222 GHRQKRVLLVVLAFITCMLPPHVRITV---INTEDSHMYFSQYFNIALQFLYLSA 348
 DB 257 NHRQVVKMLAVVVFAPILCMLPFHVRGYFSKSLRPSVAIAISQYCNLVSVFLYLSA 316
 QY 349 SINPLIYNISKYVAAKLLARSRPGRFHSRDTAGEVADGCGDVGVTETSAN 407
 DB 317 AINPLIYNISKYVAAKLLARSRPGRFHSRDTAGEVADGCGDVGVTETSAN 365
 RESULT 3
 GHSR HUMAN STANDARD; PRT; 366 AA.
 AC Q92847; Q92848; Q96R07;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing
 peptide receptor) (GHRP) (ghrelin receptor).
 GN GHSR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
 RC TISSUE=Pituitary;
 RX MEDLINE=9633798; PubMed=8688086;
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,
 RA Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
 RA Palyha O.C., Anderson J., Pares S.S., Diaz C., Chou M., Liu K.K.,
 RA McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevics M.,
 RA Pecheur A.A., Nargund R., Srinatchsingh D.J.S., Dean D.C., Meilho D.G.,
 RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
 RA "A receptor in pituitary and hypothalamus that functions in growth
 hormone release.";
 RT Science 273:974-977 (1996).
 RL [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
 RX MEDLINE=21255649; PubMed=11356716;
 RA Peterseim S., Rasch A.C., Penschorn M., Beil F.U., Schulte H.M.;

RT "Genomic structure and transcriptional regulation of the human growth
 hormone secretagogue receptor.";
 RT Endocrinology 142:2649-2659 (2001).
 RL [3]
 RN FUNCTION.
 RP MEDLINE=21219832; PubMed=11322507;
 RA Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,
 RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
 RT "Growth hormone secretagogue receptor family members and ligands.";
 RL Endocrine 14:9-14 (2001).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=20067959; PubMed=10604470;
 RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
 RT "Ghrelin is a growth-hormone-releasing acylated peptide from
 stomach.";
 RL Nature 402:656-660 (1999).
 CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
 CC Stimulates growth hormone secretion. Binds also other growth
 CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
 CC as well as non-peptide, low molecular weight secretagogues (e.g.
 CC L-692,429, MK-0677, adenosine).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
 CC SECRETAGOGUES.
 CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC
 DR EMBL; U60179; AAC50653.1; -;
 DR EMBL; U60181; AAC50654.1; -;
 DR EMBL; AF369786; AK71539.1; -;
 DR EMBL; AF369786; AK71540.1; -;
 DR Genew; HGNC:4267; GHSR.
 DR MIM; 601898; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECP_F1_1.
 DR PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Alternative splicing.
 FT DOMAIN 1 40
 FT TRANSMEM 41 66
 FT DOMAIN 67 72
 FT TRANSMEM 73 96
 FT DOMAIN 97 117
 FT TRANSMEM 118 139
 FT DOMAIN 140 162
 FT TRANSMEM 163 183
 FT DOMAIN 184 211
 FT TRANSMEM 212 235
 FT DOMAIN 236 263
 FT TRANSMEM 264 285
 FT DOMAIN 286 302
 FT TRANSMEM 303 326
 FT DOMAIN 327 366
 FT DISULFID 116 198
 FT CARBOHYD 13 13
 FT CARBOHYD 27 27
 FT VARSPLIC 266 289
 FT VARSPLIC 290 366
 SQ SEQUENCE 366 AA, 41328 MW, D1B62710DA9DC0C6 CRC64;

Query Match 39.8%; Score 857.5; DB 1; Length 366;
 Best Local Similarity 44.5%; Pred. No. 1.9e-48;
 Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

5 WNGSGPGRG-----REPWPALPCD---ERRCSPPLGALVPTAVCLFVVGSGN 56
 2 WNAAT--PSEEPNVTLLDLDMDASPGNDSPDLLPLFPAPLAGVATCVAFVVGISG 61
 57 VVTVMILIGRDMRTTNTLYLGSMVSDLLILGLPFDLRLMSRPVWFGPLCRSLY 116
 62 LITMLVSRFRRLRTTNTLYLSMAFSDLLIFLCPDLVRLMQRPVWFGDLCKLFGF 121
 117 VGEGCTYATLLMTALSVRYALICRPLARVLRRLRALIVLMAVALSAGPFLF 176
 122 VSESCTYATVLTITLALSVRYALICRPLAKVVTGRVAVIYVMAVAFSAGIFVL 181
 177 VGEVODPGISVVPGLNGTARIASSPLASPPMLSRAPPPSPGDETAALAFSREC 236
 182 VGEVHE-----NGT-----DP--W-----DINECR 199
 237 PS--PAQLGALRVMLVTTAYFFLPCLSLIYGLIGRELMSRRLRGPASGRERGR 294
 200 PEFARNSGLITVMVWVSSIFFLPVFCITVLSLIGRKLRRRRGDVAVGASLRDQNHK 259
 295 QTRKVLVAVVLAFLICMLPFRVGRITVINTEDS---RMVYFSQYENIVLQLFYLSASIN 351
 260 QTVKMLAVVVFARILCMLPFRVGRILFSKSFEGSLIAQISQYCNLVSVLFLYSALIN 319
 352 PITYNLISKYRAAFPLILAKRSRPRGFRSRDTAGVAGDGTGVGTETSAN 407
 320 PITYNLISKYRAVAVRLLGFEPFSQKSLTLDSESR-----AVTESIN 365

RESULT 4
 GHSR RAT STANDARD; PRT; 364 AA.

AC 008725;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).
 GN GHSR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SOURCE FROM N.A.
 RC TISSUE=Pituitary;
 RA MEDLINE=97246555; PubMed=9092793;
 RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C.P., Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;
 RA "Molecular analysis of rat pituitary and hypothalamic growth hormone secretagogue receptors";
 RT Mol. Endocrinol. 11:415-423 (1997).
 RN [2]
 RP SEQUENCE OF 1-240 FROM N.A.
 RC STRAIN=Wistar; TISSUE=Pituitary;
 RA MEDLINE=98100386; PubMed=9437732;
 RA Yokoe R., Sato M., Matsubara S., Ohye H., Nimi M., Murao K., Takahara J.;
 RA "Molecular cloning and gene expression of growth hormone-releasing peptide receptor in rat tissues";
 RT Peptides 19:15-20 (1998).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=20067959; PubMed=10604470;
 RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
 RA "Ghrelin is a growth-hormone-releasing acylated peptide from stomach.";
 RT Nature 402:656-660 (1999).
 CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.

CC Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine) (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U94321; AAC5156.1; -;
 DR EMBL; AB001962; BAB1777.1; ALT_INIT.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 40
 FT TRANSMEM 41 66
 FT DOMAIN 67 72
 FT TRANSMEM 73 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 139
 FT TRANSMEM 140 162
 FT DOMAIN 163 183
 FT TRANSMEM 184 211
 FT TRANSMEM 212 235
 FT TRANSMEM 236 263
 FT TRANSMEM 264 285
 FT DOMAIN 286 302
 FT TRANSMEM 303 326
 FT DOMAIN 327 364
 FT DISULFID 115 197
 FT CARBOHYD 13 13
 FT CARBOHYD 26 26
 SQ SEQUENCE 364 AA; 40963 MW; DCBF55BEO61EBE9 CRC64;

Query Match 39.7%; Score 855.5; DB 1; Length 364;
 Best Local Similarity 48.0%; Pred. No. 2.5e-48;
 Matches 184; Conservative 52; Mismatches 82; Indels 65; Gaps 11;

5 WNGSDPGRGARP-----PWPALPCD---ERRCSPPLGALVPTAVCLFVVGSG 55
 2 WNAAT--PSEEPNVTLLDLDMDASPGNDSPDLLPLFPAPLAGVATCVAFVVGISG 59
 56 NVVTVMILIGRDMRTTNTLYLGSMVSDLLILGLPFDLRLMSRPVWFGPLCRSL 115
 60 NLTMLVSRFRRLRTTNTLYLSMAFSDLLIFLCPDLVRLMQRPVWFGDLCKLFGF 119
 116 VGEGCTYATLLMTALSVRYALICRPLARVLRRLRALIVLMAVALSAGPFLF 175
 120 FVSESCTYATVLTITLALSVRYALICRPLAKVVTGRVAVIYVMAVAFSAGIFV 179
 176 VGEVODPGISVVPGLNGTARIASSPLASPPMLSRAPPPSPGDETAALAFSREC 235
 180 LVGEVHE-----NGT-----DPRD-----TNEC 197
 236 RPS--PAQLGALRVMLVTTAYFFLPCLSLIYGLIGRELMSRRLRGPASG---RE 290
 198 RATEFAVRNSGLITVMVWVSSIFFLPVFCITVLSLIGRKLW---RR--RDAAVAGASLRD 253
 291 RGHQOTKRVLLVAVVLAFLICMLPFRVGRITVINTEDS---RMVYFSQYENIVLQLFYLS 347
 254 QNHQOTKRVMLAVVVFARILCMLPFRVGRILFSKSFEGSLIAQISQYCNLVSVLFLYS 313
 348 ASINPIYNLISKYRAAFPL 370

Db 314 AAINPILVNMKRRVAVFKLL 336

RESULT 5
GHSR_MOUSE
ID GHSR_MOUSE STANDARD; PRT; 257 AA.
AC Q99P50; Q91282;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor) (Fragment).
GN GHR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-183 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RA Kacsoh B.;
RT "Cloning of mouse ghrelin/growth hormone secretagogue receptor cDNA by rapid amplification of cDNA ends (RACE)."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 73-257 FROM N.A.
RC STRAIN=129S3/SvImJ;
RA Peng X., Frohman L.A., Kineman R.D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine) (by similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AY056474; AAL1336.1; -
CC EMBL: AF332997; AAC61141.1; -
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC Pfam: PF00001; 7tm.1; 1.
CC PRINTS: PR00237; GPCRHHODPSN.
CC PROSITE: PS00237; G-PROTEIN RECP_F1_1; 1.
CC PROSITE: PS50262; G-PROTEIN RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1
FT DOMAIN 1 35
FT TRANSSEM 36 58
FT DOMAIN 59 77
FT TRANSSEM 78 100
FT DOMAIN 101 125
FT TRANSSEM 126 148
FT DOMAIN 149 160
FT TRANSSEM 161 183
FT DOMAIN 184 211
FT TRANSSEM 212 234
FT DOMAIN 235 257
FT DISULFID 115 197
FT CARBOHYD 13 13
FT CARBOHYD 26 26
FT CARBOHYD 187 187
FT NON TER 257
SQ SEQUENCE 257 AA; 28743 MW; D6FA21EAB0E30791 CRC64;
Query Match 28.5%; Score 614; DB 1; Length 257;
Best Local Similarity 45.1%; Pred. No. 5.7e-33;

Matches 137; Conservative 39; Mismatches 66; Indels 62; Gaps 10;

QY 5 WNGSDGEGAREP-----PWPALPCD---ERRSPPLGLVVTAVCLLPVGVG 55
Db 2 WNAI--PSEEPENVTLDLDWDASPGNDSDELPLPAPILAGVATCAVAFVGGIS 59
QY 56 NVVTMLIGRYDMRTTNLYLSMAVSDLLILGLPFDLYRLMSPVFPVPLCRSL 115
Db 60 NLTLMLVSRRELRTTNLYLSMAFSDLLIFLCPDLYRLMSPVFPVPLCRSL 119
QY 116 YVGGCTYATLHNTALSVRYLAICRPLRVLTTRRRVALLAVMAVALLSGPFL 175
Db 120 FVSECTYATVLTALSVRYLAICRPLRVLTTRRRVALLAVMAVALLSGPFL 179
QY 176 LVGEOPDGISVPLNGLTARLASSPLASPPLMSRAPSPSPGPTAAALFSPRC 235
Db 180 LVGEHE-----NGT-----DPRD-----TWEC 197
QY 236 RPS--PAQGLRVMLWTTAVFPLPFLCLSLYGLIGRELMSRRPLRGPAAG--RE 290
Db 198 RATEFAVRSGLLTVMWVSSVFPLPVCLTVLSLRKLM--RR--RGDAVAGSSLRD 253
QY 291 RGRH 294
Db 254 QNHK 257

RESULT 6
NTR1_RAT
ID NTR1_RAT STANDARD; PRT; 424 AA.
AC P20789;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Neurotensin receptor type 1 (NTR-1) (High-affinity leucobastine-insensitive neurotensin receptor) (NTRH).
GN NTR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90297956; PubMed=1694443;
RA Tanaka K., Masu M., Nakanishi S.;
RT "Structure and functional expression of the cloned rat neurotensin receptor."
RL Neuron 4:847-854 (1990).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC HIGHEST TO TACHYKININS RECEPTORS.
CC PIR: JH0164; JH0164.
CC InterPro: IPR000276; GPCR_Rhodopsn.
CC InterPro: IPR003985; NTL_Rec.
CC InterPro: IPR003984; NT_Rec.
CC Pfam: PF00001; 7tm.1; 1.
CC PRINTS: PR00237; GPCRHHODPSN.
CC PRINTS: PR01479; NEUROTENSINR.
CC PROSITE: PS00237; G-PROTEIN RECP_F1_1; 1.
CC PROSITE: PS50262; G-PROTEIN RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate.
KW DOMAIN 1
FT DOMAIN 1 64
FT TRANSSEM 65 87
FT DOMAIN 88 96
FT TRANSSEM 97 121
FT DOMAIN 122 143
FT TRANSSEM 144 165

FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 189 210 4 (POTENTIAL).
 FT DOMAIN 211 235 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 236 260 5 (POTENTIAL).
 FT DOMAIN 261 308 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 309 330 6 (POTENTIAL).
 FT DOMAIN 331 348 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 349 372 7 (POTENTIAL).
 FT DOMAIN 373 424 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 142 225 BY SIMILARITY.
 FT LIPID 388 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA, 47054 MW, A9C2F7A9B9BCD3 CRC64;

Query Match 23.1%; Score 498; DB 1; Length 424;
 Best Local Similarity 32.9%; Pred. No. 2.5e-25;
 Matches 125; Conservative 58; Mismatches 111; Indels 86; Gaps 11;

QY 39 VPTAVCLCLFVVGSGNVVTMLIGR---YRDMRTTNLYLGSMAVSDLLI-LGLPFD 94
 DB 65 VLTAVLYALFVVGTVGNSVTAFTLARKKSLQSLQSTVHVLGSLSDLLILLAMPVE 124
 QY 95 LVR-LMRSRPWVFGPLCLSLVVGECYATLHMTALSVERYLAIICRPLARAVLTR 153
 DB 125 LYNFTVWHHPMAFGDAGCGRYFELRDACYATALNVASLSVERLAIICHPKAKTMSRS 184
 QY 154 RYRALIAVMAVALLSAGPFLFVVGEDPGISVVPGLNGTARIASSPPLASPLMLSRA 213
 DB 185 RTKKFISATWLASALLAIPLFTFTGLQNRSGDGTGPG---GLVCTPIVDY 231
 QY 214 PPPSPSGPETAALALFSRECRPSAOLGALRYMLVTTAYFPL-PFLCLSLIYGLIGR 272
 DB 232 -----ATVKKVIVQNTFMSFLPMLVSIINTVIAN 262
 QY 273 ELWSSRRPLRGPASGR-----ERGHQTR-----VLTAVVLAFTI 309
 DB 263 KLTVM---VHQAAEQGVCTVGTNGLEHSTFNMTIEGRVQALRHGVLAIVAVV 319
 QY 310 CWLRFVHGRITVINTEDSR---MMYFSQYFNIVALOLFYLASINPILVNLISKYRAA 365
 DB 320 CWLRFVHGRITVINTEDSR---MMYFSQYFNIVALOLFYLASINPILVNLISKYRAA 365
 QY 366 AFKLLA-----KRSRP 377
 DB 380 FLSTLACLPGRHRRKKRP 399

RESULT 7
 NTRL MOUSE STANDARD; PRT; 424 AA.
 AC 088319;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotensin receptor type 1 (NT-R-1).
 GN NTSR1 OR NTSR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Snider J., Sano H., Ohta M.;
 RT "Neurotensin receptor type 1,"
 RT Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RL - FUNCTION: RECEPTOR FOR THE TRIDECAPPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

HIGHEST TO TACHYKININS RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB017027; BAB3013.1; -
 DR MGD; MG197386; Nser.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR003985; NT1_rec.
 DR InterPro; IPR003984; NT1_rec.
 DR Pfam; PF00001; 7tm.1; 1-
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PRINTS; PR01479; NEUROTENSINR.
 DR PRINTS; PR01480; NEUROTENSINR.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS50262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 65 87 1 (POTENTIAL).
 FT DOMAIN 88 96 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 97 121 2 (POTENTIAL).
 FT DOMAIN 122 143 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 144 165 3 (POTENTIAL).
 FT DOMAIN 166 188 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 189 210 4 (POTENTIAL).
 FT DOMAIN 211 235 5 (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 308 7 (POTENTIAL).
 FT TRANSMEM 309 330 8 (POTENTIAL).
 FT DOMAIN 331 348 9 (POTENTIAL).
 FT TRANSMEM 349 372 10 (POTENTIAL).
 FT DOMAIN 373 424 11 (POTENTIAL).
 FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 141 224 BY SIMILARITY.
 FT LIPID 388 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA, 47216 MW, 889A73171A48711 CRC64;

Query Match 23.0%; Score 495.5; DB 1; Length 424;
 Best Local Similarity 33.6%; Pred. No. 3.6e-25;
 Matches 127; Conservative 56; Mismatches 114; Indels 81; Gaps 12;

QY 39 VPTAVCLCLFVVGSGNVVTMLIGR---YRDMRTTNLYLGSMAVSDLLI-LGLPFD 94
 DB 64 VLTAVLYALFVVGTVGNSVTAFTLARKKSLQSLQSTVHVLGSLSDLLILLAMPVE 123
 QY 95 LVR-LMRSRPWVFGPLCLSLVVGECYATLHMTALSVERYLAIICRPLARAVLTR 153
 DB 124 LYNFTVWHHPMAFGDAGCGYFELRDACYATALNVASLSVERLAIICHPKAKTMSRS 183
 QY 154 RYRALIAVMAVALLSAGPFLFVVGEDPGISVVPGLNGTARIASSPPLASPLMLSRA 213
 DB 184 RTKKFISATWLASALLAIPLFTFTGLQNRSGDGTGPG---GLVCTPIVDY 231
 QY 214 PPPSPSGPETAALALFSRECRPSAOLGALRYMLVTTAYFPL-PFLCLSLIYGLIGR 272
 DB 214 ADGQHGCG-----LVCTPT-VDTATVAVVIVQNTFMSFLPMLIISINTVIAN 261
 QY 273 EL---WSSRRPLRGPASGR-----ERGHQTR-----VLTAVVLAFTI 311
 DB 262 KLTVMHQAABEQRGVCTVGTNGLEHSTFNMTIEGRVQALRHGVLAIVAVVAVV 321
 QY 312 LRFVHGRITVINTEDSR---MMYFSQYFNIVALOLFYLASINPILVNLISKYRAAF 367
 DB 322 LRFVHGRITVINTEDSR---MMYFSQYFNIVALOLFYLASINPILVNLISKYRAAF 367

OY 368 KLLA-----RKSRP 377
DB 382 STIACLCGWRRRRRKKRP 399

RESULT 8

NTRI_HUMAN STANDARD; PRT; 418 AA.
AC P30989; Q9H4T5; Q9H4H1;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucobactine-
sensitive neurotensin receptor) (NTRH).
OS NTR1 OR NTR.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93154505; PubMed=8381365;
RA Vita N., Laurent P., Lefort S., Chalon P., Dumont X., Kaghad M.,
RA Gully D., Le Fur G., Ferrara F., Caput D.;
RT "Cloning and expression of a complementary DNA encoding a high
RT affinity human neurotensin receptor.";
RL FEBS Lett. 317:139-142(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill M.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clegg S.,
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A.G., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Garner P.,
RA Grafham D.V., Griffiths J.A., Fraser A., French L., Garner P.,
RA Hammon S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levenshlo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McComachie L.J., McLeay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehara H.K., Showkneen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,
RA Swann R.M., Symcote N., Taylor R., Tee I., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.W.,
RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X70070; CAA49675.1; -
DR EMBL; AL357033; CAC14923.1; -
DR EMBL; AL035669; CAC12747.1; -
DR PIR; S29506; S29506.
DR HSP; P02699; 1F88.
DR Genew; HGNC:8039; NTR1.
DR MIM; 162651; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR003985; NT1_Rec.
DR InterPro; IPR003984; NT_Rec.
DR Pfam; PF00001; 7tm_1; 2_Fec.
DR PRINTS; PRO1479; NEUROTENSINR.
DR PRINTS; PRO1480; NEUROTENSINR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1.
DR PROSITE; PS00262; G_PROTEIN_REC_P2_1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 63
FT TRANSMEM 64 86
FT TRANSMEM 87 95
FT TRANSMEM 96 120
FT DOMAIN 121 142
FT TRANSMEM 143 164
FT DOMAIN 165 187
FT TRANSMEM 188 209
FT DOMAIN 210 234
FT TRANSMEM 235 259
FT DOMAIN 260 303
FT TRANSMEM 304 325
FT TRANSMEM 326 343
FT TRANSMEM 344 367
FT DOMAIN 368 418
FT CARBOHYD 4 4
FT CARBOHYD 37 37
FT CARBOHYD 41 41
FT DISULFID 141 224
FT LIPID 383 383
FT CONFLICT 200 200
SQ SEQUENCE 418 AA; 46288 MW; BBBDECB2B8E390 CRC64;

Query Match 22.8%; Score 491; DB 1; Length 418;
Best Local Similarity 32.8%; Pred. No. 6; 9e-25;
Matches 125; Conservative 62; Mismatches 118; Indels 76; Gaps 13;

OY 39 VPPTAVCLCFVGVGSGNVTVNLIGR---YRDRRTTNLYLSMAVSDIL-IILGPF 94
DB 64 VLVTVAVYALFVGVGVGNVTATFTLARKKSLQSTVHVLGLSLDLSLTLMPVE 123
OY 95 LYR-LWRSRPVFGPPLCRSLVVGECYATLHMTALSEVRYLAIQPLRARVLVTR 153
DB 124 LVNFIWVHHWAFGADGCRGVYFLRDACTATLNLVNSLSEVRYLAIQPLKATIMRS 183
OY 154 RVVALAVLMAVALLSAGPPLFLVGVHODGISVPELNGTARIASSPLASSPPLWSRA 213
DB 184 RTKFFISAIWLSALLTVPLFTMG-EQN----- 211
OY 214 PRSPSPSGPETAALALFSECRSPAPQALRVLMTVT-AVFLEFLCLISLYGLIGR 272
DB 212 -----RSAGQAGAGLVCTPT-IHTATVKVIOVNTFMGFIIPMVVISTVNTIAN 261
OY 273 ELWSSRRPL-RGPAS-----GRGRHQRKRVLLVVLVLAFLICWLPFHV 316
DB 262 KLTVMVROAAFGQOVCTVGGESHSTFMAIRPGRVQALRHGVRLRAVVAVFVWCWLPYHV 321
OY 317 GRIIVINTEDSR---MMY-FSQFNFVVALQFLYLSASINPLVNLISKRY-----AA 365
DB 322 RRLMFCYISDEQWTPPLDYFYHYFVWTVNALFYVSSINPLVNLVSANRHFILATLAC 381
OY 366 AFKLLARKSRPGFHRSDT 386
DB 382 LCPVWRKKRP-AFSRKADS 401

Matches	129	Conservative	60;	Mismatches	135;	Indels	90;	Gaps	16;
QY	17	PPPPALP--	-CDERCSPPGLALVPPNVAVCLCFVGVSGNVVVVMI-----	GR	65				
Db	8	PPPPSPNGSLSDARGLVDTRLMKAVLFALVALWALGAGNALSVHYLKRARAGR		67					
QY	66	YRMRRTTNLYIGSMVSDLLIL--GLPFDLYR-LWRSPWVGGPLLCRLSLYVGGCTY		123					
Db	68	LRH-----HVLSTLAGLILLVGVPEVLSFWFPHPWVGDLGCGGYVVEHLCAV		120					
QY	124	ATLLHMTALSVERYLAICRLPRARVLYTRRRVALIAVIMANVALSAGPLFLVGV----	179						
Db	121	ATVLSVAGLSAEKCLAVCQPLRARSILTRRTKWLVALSWASLGLALPMAVIMQKHKL	180						
QY	180	----EODPELSVPGNGNRIASS-----PLASSPL-----	MTS	211					
Db	181	ETADGEPEPSRRCVYL--VSRITALQVFIQVNVLSFVPLATATLNGVTVSHLLALS	238						
QY	212	RAPPSPSGPEFVAEAAALFSPRECRPSPAQALRYMLWTTAYFPLPLCLSLVGLIG	271						
Db	239	QVBSSTFP-GSSTSPRLLESEE-----GLSLFVWKKT-----	FIG--G	276					
QY	272	RELWSSRRPLRGPASGRGHRQTKRVLLVTVLAFICMLPEHVGRILYINEDSR---	328						
Db	277	QVSLVRKQVR-----RISLRQSVQVLEAIWVYICWLPHARLMCYVPDAMTD	330						
QY	329	MMFSGYFNIVALQLFYLSASINPLIYNIISKRYAAKLLARKSRGRPH	381						
Db	331	PLVNFYHYFMVNTTIFYVSSAVTPPLLYNAVSSSF--KLFLVAVSLGGEH	380						
RESULT	10								
NTR2	MOUSE	STANDARD;	PRT;	417	AA.				
ID	NTR2_MOUSE								
AC	P70310;								
DT	01-NOV-1997 (Rel. 35, Created)								
DT	01-NOV-1997 (Rel. 35, last sequence update)								
DT	16-OCT-2001 (Rel. 40, last annotation update)								
DE	Neurotensin receptor type 2 (NT-R-2) (low-affinity leucobastine-sensitive neurotensin receptor) (NTRL).								
DE	NTSR2.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.								
CX	NCBI_TaxId=10090;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=BALB/c; TISSUE=Brain;								
RX	MEDLINE=96388216; PubMed=8795617;								
RA	Medella J., Botto J.-M., Guillemare E., Coppola T., Sarret P.,								
RA	Vincent J.-P.,								
RT	"Structure, functional expression, and cerebral localization of the								
RT	leucobastine-sensitive neurotensin/neuromedin N receptor from mouse								
RL	brain."								
CC	J. Neurosci. 16:5613-5620(1996).								
CC	-1- FUNCTION. RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS								
CC	ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-								
CC	CALCIUM SECOND MESSENGER SYSTEM.								
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.								
CC	-1- TISSUE SPECIFICITY: EXPRESSED MAXIMALLY IN THE CEREBELLUM,								
CC	HYPOCAMPUS, PIRIFORM CORTEX AND NEOCORTEX OF ADULT BRAIN.								
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED POORLY IN 7-DAY-OLD BRAIN.								
CC	EXPRESSION INCREASES AT DAY 15 TO REACH A MAXIMAL LEVEL IN 35-DAY-								
CC	OLD BRAIN.								
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.								
CC	HIGHEST TO TACHYKININS RECEPTORS.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute								

or send an email to license@sb-sib.ch).

CC -----
 CC EMBL; U51908; AAB17285.1; -
 DR MGD; MGI:108018; Ntsr2.
 DR InterPro; IPR000276; GPCR Rhodopsin.
 DR InterPro; IPR003986; NT2 Rec.
 DR InterPro; IPR003984; NT Rec.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PRO1479; NEUROTENSINR.
 DR PRINTS; PRO1481; NEUROTENSINR.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
 KM G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1
 FT TRANSMEM 33 55
 FT DOMAIN 56 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 176
 FT DOMAIN 177 217
 FT TRANSMEM 218 238
 FT DOMAIN 239 298
 FT TRANSMEM 299 319
 FT DOMAIN 320 338
 FT TRANSMEM 339 359
 FT DOMAIN 360 417
 FT DISULFID 108 194
 FT LIPID 378 378
 SQ SEQUENCE 417 AA; 46537 MW; EBRDDB6507223DD CRC64;

Query Match 18.7%; Score 404; DB 1; Length 417;
 Best Local Similarity 30.6%; Pred. No. 2.6e-19;
 Matches 126; Conservative 58; Mismatches 146; Indels 82; Gaps 14;

QY 19 WPALP-----PCDERRCSPPLGALVPTAVCLCFVVGSGNVVTWMLIGRYDMRT- 71
 Db 7 WPPRPSAGSLSEARLGVDTRLMAKVFTALYSILFAGTGNALSHVHLKARTGRPG 66
 QY 72 TTNYLGSMVSDLLIL-GLPFDLYR-LMRSRPWFPGPLCRLSLYVGGCTYATLLH 129
 Db 67 RLRYHVLSTLALLLLISVMEIYFWSHVPWFGLGCRGYFVVELCAVAVLSV 126
 QY 130 TALSVETLALCRPLRAVLTTRVRALLAVLWVALISAGPFLVGVGDDPGISVVP 189
 Db 127 ASLSAERCLAVCOPLRARRLTPRRCTRLSLVWVASLGALPMAYIMQKHE----- 179
 QY 190 GLNGTARIASPLASPPMLMSRAP-PPSPSGPTAEALALFSPRCRSPALQ---GAL 245
 Db 180 -----MERADGEPFAS-----RVCTVIVSRASSRSTQVKRAGLL 215
 QY 246 RVMLVMTTA-----YFPLPFLC-----LSIL--YGLIGREIMSSRPL 281
 Db 216 RSPIMELTALINGITVNLVALYSGVPSASQVNSIPSLLELISEGGLGFTWKRTLSL 275
 QY 282 RGPASAGREGRHQTK-----RVLLVVVLATTCMLPFPVGRITTYNTD---SRMY 331
 Db 276 GVQASLVNHHKASQTRLSQHSAGVLRALVAVVVICMLPYHARLWVCYLPDGGWTDLYD 335
 QY 332 FSOYFNIVIALQFLYLSASINPLLYLISKYRAAPFKILLARKSRPRGRHS 383
 Db 336 FYHIFVWVNTLTFVYSSAVTVLVNAVSSFR---KLFLESLSLGEGORS 383

RESULT 11
 NTR2 RAT
 ID NTR2 RAT STANDARD; PRT; 416 AA.
 AC 063384;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurotensin receptor type 2 (NT-R-2) (High-affinity leucocabastine-

DE sensitive neurotensin receptor).

GN NTR2 OR NTR2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Hypothalamus;

RX MEDLINE=96228041; PubMed=8647296;

RA Chalon P., Vita N., Kaghad M., Guillemont M., Bonin J.,

RA Delpech B., Le Fur G., Ferrara P., Caput D.;

RT "Molecular cloning of a leucocabastine-sensitive neurotensin binding

RT site.";

RT FEBS Lett. 386:91-94(1996).

-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS

ASSOCIATED WITH G-PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-

CALCIUM SECOND MESSENGER SYSTEM.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER

LEVELS SEEN IN THE HEART AND INTESTINE.

-1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND

EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD

BRAIN).

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

HIGHEST TO TACHYKININS RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.sib-sib.ch/announce/>

or send an email to license@sb-sib.ch).

CC -----
 CC EMBL; X97121; CA65787.1; -

DR InterPro; IPR000276; GPCR Rhodopsin.

DR InterPro; IPR003986; NT2 Rec.

DR InterPro; IPR003984; NT Rec.

DR Pfam; PF00001; 7tm_1; 1.

DR PRINTS; PRO1479; NEUROTENSINR.

DR PRINTS; PRO1481; NEUROTENSINR.

DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.

DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.

KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.

FT DOMAIN 1 32
 FT TRANSMEM 33 55
 FT DOMAIN 56 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 176
 FT DOMAIN 177 217
 FT TRANSMEM 218 238
 FT DOMAIN 239 298
 FT TRANSMEM 299 319
 FT DOMAIN 320 338
 FT TRANSMEM 339 359
 FT DOMAIN 360 417
 FT DISULFID 108 194
 FT LIPID 377 377
 SQ SEQUENCE 416 AA; 46265 MW; 127FC5F5CB6FE208 CRC64;

Query Match 18.6%; Score 401; DB 1; Length 416;
 Best Local Similarity 29.4%; Pred. No. 4.1e-19;
 Matches 122; Conservative 68; Mismatches 145; Indels 80; Gaps 13;

QY 18 WPALP-----PCDERRCSPPLGALVPTAVCLCFVVGSGNVVTWMLIGRYDMRT 71
 Db 6 WPPRPSAGSLSEARLGVDTRLMAKVFTALYSILFAGTGNALSHVHLKARTGRPG 65
 QY 72 TTNYLGSMVSDLLIL-GLPFDLYR-LMRSRPWFPGPLCRLSLYVGGCTYATLLH 128

DT 01-JUN-1993 (Rel. 26, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Somatostatin receptor type 5 (SSSR).
 GN SSTR5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=93125493; PubMed=1362243;
 RA O'Carroll A.-M., Lolait S.J., Korig M., Mahan L.C.;
 RT "Molecular cloning and expression of a pituitary somatostatin
 RT receptor with preferential affinity for somatostatin-28.";
 RL Mol. Pharmacol. 42:939-946(1992).
 RN [2]
 RP REVISIONS TO C-TERMINUS.
 RC TISSUE=pituitary;
 RX MEDLINE=94195267; PubMed=7908405;
 RA Panetta R., Greenwood M.T., Warszynska A., Demchyshyn L.L., Day R.,
 RA Niznik H.B., Srikant C.B., Patel Y.C.;
 RT "Molecular cloning, functional characterization, and chromosomal
 RT localization of a human somatostatin receptor (somatostatin receptor
 RT type 5) with preferential affinity for somatostatin-28.";
 RL Mol. Pharmacol. 45:417-427(1994).
 CC -1- FUNCTION: RECEPTOR FOR SOMATOSTATIN-28. THE ACTIVITY OF THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBIT ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: PROMINENT IN THE PITUITARY AND SMALL
 CC INTESTINE. LOW LEVELS IN ISLETS AND SPLEEN. NOT DETECTED IN
 CC KIDNEY, PANCREAS, CEREBELLUM, OR CORTEX.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; L04535; AAA17029.1; -;
 DR EMBL; U01152; AAC09011.1; -;
 DR EMBL; X74828; CAAS2825.1; -;
 DR HSSP; P02699; 1F88.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1.1;
 DR PRINTS; PR00237; GPCRHOODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
 DR PROSITE; PS50262; G-PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate.
 FT DOMAIN 1 35
 FT TRANSMEM 36 63
 FT TRANSMEM 64 73
 FT TRANSMEM 74 99
 FT DOMAIN 100 111
 FT TRANSMEM 112 133
 FT DOMAIN 134 155
 FT TRANSMEM 156 176
 FT DOMAIN 177 196
 FT TRANSMEM 197 221
 FT DOMAIN 222 247
 FT TRANSMEM 248 273
 FT DOMAIN 274 283
 FT TRANSMEM 284 308
 FT DOMAIN 309 363
 FT TRANSMEM 364 383
 FT CARBOHYD 13 13
 FT CARBOHYD 23 23
 FT CARBOHYD 186 186
 FT DISULFID 110 185

FT LIPID 320 320 PALMITATE (POTENTIAL).
 SQ SEQUENCE 363 AA; 39971 MW; 4BD4512960613B4A CRC64;
 Query Match 16.6%; Score 358; DB 1; Length 363;
 Best Local Similarity 27.3%; Pred. No. 2.1e-16;
 Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;
 QY 5 WNGSDGEGAEPEPMPALPCDEKRCSPFLGA---LYPVAVCLFLVGVGVVVTM 61
 DB 12 WNASAASSGNHN--WSLVG-----SASPWGAARVLPVLYLVC--TVGLSGNTLVY 60
 QY 62 LIGERYDMRTTNLYIGSMAYSDLLILGLPDIYRLWRSRWVGPPLCRLSLVYEGC 121
 DB 61 VVLRHAKKKTNYVYIINLAADVLFMLGLPFLATQNAVGVWPGSLCRLVMTLDGIN 120
 QY 122 TYALLMTALSVRYALICPLPARVLYTRRRBALIYVMAVALSAGPLFLVGEQ 181
 DB 121 QFTSIFCLMWASVDRYLVAVHPPLRSARWRBRVAKMASAAVVFSLMSLPILVFDVQE 180
 QY 182 DPGISVPLNGTARIASPLASSPPLWLSRAPPPSPSGPETAALAFSECRPSPAQ 241
 DB 181 G-----WGTONLS-----W-----PEPVG 194
 QY 242 LGLRVMWLTAVFPLPLCLSLYGLI-----GRELWSSRRPLRGPAASGRGRGR 294
 DB 195 LMGAAFTYTSVIGFPGPLPLVICCYLLIVKVKAKGRVQSSR-----RRRSEP 244
 QY 295 QTRKVLVLAFLICPLPFHVGIIYIN--TEDSRMYFSGYFNIVLQLYLSSTPI 353
 DB 245 KVRMVAVVVLPVPGWLPFFIVNVLAFPLPEEPTSAGLYFVVVLS--YANSCANPL 302
 QY 354 LVNLSKKYRAAFKLLARKSRPGRHRSRDTAGEVAGDTGG 396
 DB 303 LYGLISDNFQSPFRKVLCLR---RGYGMEDADAIFERPDSG 341
 RESULT 14
 ID SSR3_HUMAN STANDARD; PRT; 418 AA.
 AC P32745;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Somatostatin receptor type 3 (SSSR) (SSR-28).
 GN SSTR3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93149123; PubMed=1337145;
 RA Yamada Y., Reisine T., Iaw S.F., Iliara Y., Kubota A., Kagimoto S.,
 RA Saito M., Saito Y., Bell G.I., Saito S.;
 RT "Somatostatin receptors, an expanding gene family: cloning and
 RT functional characterization of human SSTR3, a protein coupled to
 RT adenylyl cyclase.";
 RT Mol. Endocrinol. 6:2136-2142(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93338970; PubMed=8097479;
 RA Corness J.D., Demchyshyn L.L., Seeman P., van Tol H.H.M.,
 RA Srikant C.B., Kent G., Patel Y.C., Niznik H.B.;
 RT "A human somatostatin receptor (SSTR3), located on chromosome 22,
 RT displays preferential affinity for somatostatin-14 like peptides.";
 RL FEBS Lett. 321:279-284(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Bagdely C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.F., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,

RL Genomics 46:426-434(1997). Integral membrane protein.
 CC -1- SUBCELLULAR LOCATION: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF034633; AAC26082.1; -
 DR Genew; HGNC:4496; GPR39.
 DR MIM; 602886; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7cm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 25
 FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT DOMAIN 47 71
 FT TRANSMEM 72 92
 FT DOMAIN 93 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 175
 FT DOMAIN 176 223
 FT TRANSMEM 224 243
 FT DOMAIN 244 280
 FT TRANSMEM 281 301
 FT DOMAIN 302 321
 FT TRANSMEM 322 347
 FT DOMAIN 348 428
 FT DISULFID 108 210
 FT CARBOHYD 192 192
 FT CARBOHYD 206 206
 FT CARBOHYD 212 212
 SQ SEQUENCE 453 AA; 51328 MW; 8E3A233420D9021E CRC64;
 Query Match 16.2%; Score 349.5; DB 1; Length 453;
 Best Local Similarity 28.5%; Pred. No. 8.8e-16; Indels 45; Gaps 13;
 Matches 111; Conservative 64; Mismatches 170;
 QY 20 PALPDCRRCS-----PPLGALVPT--AVCLCLFVVGSGNVTV---MLIGR 65
 DB 4 PSLPQSD--CSQIDHSHVPEFEVATWIKITLIVYLIIIFWGLGNATIRVQVLOK 60
 QY 66 YRDKRTTNVIGSMVAVDILL-LGLPFDLYR-LMRSPWVFGPLLCLRLSLYVGECTY 123
 DB 61 KGYLQKEVTDHNVSLACSDILVFLIGMPEFYSIINNPULTSSYTLSCKLHFLFEACSY 120
 QY 124 ATLLHMTALSVERYLAICPLRARVLTTRRVALIAVAVALLSAGPLFLVGVEDP 183
 DB 121 ATLLHVLTLSPERYTAICHPRYKAVSGCQVYLIGFWVTSALVALPLFLAMGTEY-P 179
 QY 184 GISVVPGLNGTARLASSPLASSPPLMSRAPSPSPSGPETAEAALFSRECRPSPAQ-- 241
 DB 180 LVN-VPSHRLGLTCNRSSSTRHNEQ-----PETSMSICTNLSSRWTVFQSS 223
 QY 242 -LGLARVWLWTTAFAFLPFLCLSLIYGLIGRELMSRRP--LRGPAASGREGRHROTQR 298
 DB 224 IFGAFVYLVLLVLSAFVCMNMVQVLMKSQKSLAGGTRPQLRKSESESESTRARROTI 283
 QY 299 VLVVVLAFIICWLPFHVGRILYI--NTEDSRMYFSQYFNIVALQ--LFYLSASINPIL 354
 DB 284 FLRLIVVTLAVCMWPNQIRITMAAKPKHDWTRSYFRAYMILLPSETFFYLSVINPIL 343
 QY 355 YNLISKYTRAAAFKLLLARKSRPRGFRSR 384
 DB 344 YTVSSQQFRFVAVGVLCCLRLSLQHANHEKR 373

Search completed: March 16, 2003, 15:04:19
 Job time : 15 secs

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:10 ; Search time 18 Seconds
(without alignments)
2200.410 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156

Sequence: 1 MGSPWNGSDGEGAREPPMP.....DTGGDTVGYTETSANVKTWG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
5: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	23.1	424	2	JH0164
2	491	22.8	418	2	neurotensin recept
3	434	20.1	418	2	S29506
4	401	18.6	416	2	protein Ki0B4.4
5	378	17.5	416	2	neurotensin recept
6	358	16.6	378	2	hypothetical prote
7	356.5	16.6	363	2	somatostatin recep
8	343	15.9	363	2	somatostatin recep
9	343	15.9	363	2	somatostatin recep
10	342.5	15.9	380	2	somatostatin recep
11	338	15.7	380	2	thyrotropin releas
12	338	15.7	380	2	thyrotropin releas
13	334	15.5	380	2	thyrotropin releas
14	334	15.5	380	2	thyrotropin releas
15	333.5	15.5	367	2	somatostatin recep
16	333.5	15.5	367	2	somatostatin recep
17	333.5	15.5	367	2	somatostatin recep
18	333.5	15.5	367	2	somatostatin recep
19	330.5	15.3	428	2	G protein-coupled
20	329.5	15.3	428	2	probable G protein
21	328	15.2	380	2	somatostatin recep
22	327	15.2	380	2	orphan opioid rece
23	326.5	15.1	384	2	kappa opioid recep
24	325.5	15.1	372	2	kappa opioid recep
25	321	14.9	384	2	delta opioid recep
26	320.5	14.9	372	2	delta opioid recep
27	320	14.8	398	2	thyrotropin-relas
28	319	14.8	519	2	thyrotropin recep
29	319	14.8	372	2	delta opioid recep

30	318.5	14.8	387	2	JC5949	galanin receptor 2
31	318	14.7	380	2	I38435	angiotensin recept
32	317	14.7	398	1	JO1059	neurokinin 2 recep
33	317	14.7	411	2	I56444	thyrotropin-relea
34	317	14.7	412	2	S23436	thyrotropin-relea
35	316	14.7	393	2	A39251	thyrotropin-relea
36	314	14.6	391	2	C41795	somatostatin recep
37	311	14.4	373	2	JE0087	delta opioid recep
38	311	14.4	384	1	S00516	neurokinin 2 recep
39	310	14.4	391	2	A39297	somatostatin recep
40	308.5	14.3	369	2	D41795	somatostatin recep
41	308.5	14.3	392	2	S65693	opioid receptor mu
42	308.5	14.3	400	2	I56553	mu opiate receptor
43	308	14.3	391	2	A41795	somatostatin recep
44	308	14.3	398	2	I56517	mu opioid receptor
45	307.5	14.3	369	2	A45291	somatostatin recep

ALIGNMENTS

RESULT 1

JH0164
neurotensin receptor - rat
C/Spectrum: Rattus norvegicus (Norway rat)
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000
C/Accession: JH0164
R/Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A/Title: Structure and functional expression of the cloned rat neurotensin receptor.
A/Reference number: JH0164; PMID:90297956; PMID:1694443
A/Accession: JH0164
A/Molecule type: mRNA
A/Residues: 1-424 <TRAN>
C/Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor. The
lter (neurotensin) in the brain and as a hormone) cellular mediator in peripheral tis
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F/97-121/Domain: transmembrane #status predicted <TM1>
F/144-165/Domain: transmembrane #status predicted <TM2>
F/189-210/Domain: transmembrane #status predicted <TM3>
F/236-260/Domain: transmembrane #status predicted <TM4>
F/309-330/Domain: transmembrane #status predicted <TM5>
F/348-372/Domain: transmembrane #status predicted <TM6>
F/38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 23.1%; Score 498; DB 2; Length 424;
Matches 125; Conservative 58; Mismatches 111; Indels 86; Gaps 11;

QY	39	VPTAVCLCFVGVSGNVTVWLIGR---	YRDMRTTNLYLGSMAVSLLI-LLGLPFD 94
DB	65	VLTVAIYLLAFVGVGNSVTAFTRKKSLSQISQSVHVLGSLASDLLILLAMPVE 124	
QY	95	LVR-LMRSRPWVGPPLCRSLTLVVGECYATLHMTALVEERYLATCRPLARAVLTR 153	
DB	125	LVNFIWVHPWAGDAGCGRYVFLRDACVATATNANVSIVERTLACHPKATLMSRS 184	
QY	154	RVALAVLMAVALLSAGPFLVGVQDPGISVVGELNSTARIASSPLASSPMLMSRA 213	
DB	185	RTKRFISAIWLSALALAIPLFTMGQNSRGDTHPG-----GLVCTPIYDT----- 231	
QY	214	PPSPSPSGPETAALALFSECRPSPAQALRVMLVMTTAYFPL-PLCLISLYGLIGR 272	
DB	232	-----ATVKVIVQVNTFMSFLFPMVLVISIINTVIAN 262	
QY	273	ELWSSRRPLRGPAASGR-----ERGRQTR-----VLLVVLAFII 309	
DB	263	KLTVM---VHQAAEGRCVCTVGTGHNGLSHSTFMWTEIPKGVQALRHGVLLRAVIAFVV 319	
QY	310	CHLPFVIGRIIYINTDSR---MMTFQSYFNIVALQFLSLASINDILNLSIKYRRA 365	

RESULT 2

neurotensin receptor - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C:Accession: S29506
R:Vita, N.; Laurent, P.; Lefort, S.; Chalton, P.; Dumont, X.; Kagnad, M.; Gully, D.; le F
FEBS Lett. 317, 139-142, 1993
A>Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
A:Reference number: S29506; MUID:93154505; PMID:8381365
A:Accession: S29506
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-418 <VT>
A:Cross-References: EMBL:X70070, NID:g35020, PIDN:CAA49675.1, PID:g35021
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 22.8%; Score 491; DB 2; Length 418;
Best Local Similarity 32.8%; Pred. No. 4e-32;
Matches 125; Conservative 62; Mismatches 118; Indels 76; Gaps 13;

QY	39	VPTAVLCCTFVVGVSGNVTVTMLIGR---	VRDMRTTNLYGSMVAVSDDL-ILGLPRD	94
Db	64	VLVTAIVTALFVGTGTGNTVTAFTLARKKSLDSLOSTVHYHGLSIALSDLTITLLAMPVE	:	123
QY	95	LVR-LMRSRPWFQPLLCRLSLVVGEGCTYATLTMNLASVERVYATCRPLARVLVTR	:	153
Db	124	LVNFIVWHHPAFADAGRGYYFLDLACTYATALNVASSVSRVYLAIChDPFAKTLMSSRS	:	183
QY	154	RVRALLIIVNAVALISAGPFLFLVGEVEDPGISVVPGLNGTNRIASPLASPWLMSRA	:	213
Db	184	RTKKFISAIVWASALLTVPMFLTWG-EQN-----	:	211
QY	214	PSPSPGPETFAEAAALFSRECRPSACLGALVMMLVTT-AVFELPCLSLTYGLIGR	:	272
Db	212	-----RSADQHAGGLVCTPT-IHTATVKVVIQVMTFMSFIIPMVVISLVNIIAN	:	261
QY	273	ELWSSRRPL-RGPAA-----GREGRHQTKRVLLVLAFTICMLPEHV	:	316
Db	262	KLTGWVROAAEQGVCTVGGEHSTFSMAIEPGRVQLRHGVRLRAVIAFFVCWLPYHV	:	321
QY	317	GRIITYINTEDSR---MMT-FSOYENIVALQLFYLSAINPIILYNLISKYR-----AA	:	365
Db	322	RRIMFCYISDQMPPFLVDIFYHFYMVNALFYVSSITNPILYNLVSANFRHI FLATLAC	:	381
QY	366	AKLLARKSRPGRPHRSRD	:	386
Db	382	LCPVWRRRRRKP-AFSRKADS	:	401

RESULT 3

A:Accession: A88013
protein KiO84.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88013
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88013
A>Status: preliminary
A:Molecule type: DNA

[illegible]

```

QY      72  -TTNLYSMVSDLLIL--GLPDIYR-LWRSPNMFEGCLRLSLVGEQTYATLIL 128
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      66  GLRLRHVLSLALLSLILVSMPEWELINFWMSHPWFGDGRGYIFVRELCATVLS 125
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      129  MTLASVEERYLAICRPLRVARLVTRRRVRALAVLMAVALISAGPFLILGV----- 179
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      126  VASIASAEKCLAVCQPLRRRRRLTLTRRRRLSLVWASLGLALPMAVIMQKHEVSSAG 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      180  EDDPETSIVPGINGTARIA-----SSPIASSPPLMLSRAPSPSPSPETAEEA 228
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      186  EEPFASRYCTVLVSRATLQVFIQNVVLVSFALPLATAPL-----NGITVNHLM 234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      229  ALFSNECPSPRQGLARVMLMTTAYFPLPFLCLSLIYGLTRELWSSRPRLRCPAASG 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      235  ALYS-QVPSASAOVSSIPSRLELASE-----EGILFETWRKTLSLGVQASHV 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      289  RERGRHQRK-----RYLVLVVLAFFICMLPEHVGRIIYNTEED-----SRMYFSGEYFI 338
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      282  RHKDSAQIRSLQHSAQVRAIVAVYICWLPYHARRIMYCIIPDDGWTNELVDGYHYFYM 341
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      339  VALQIFYLASINPILVYLMLISKVYRAAFKILLARKRGPGRFHSRSDTAGEVAGD 393
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      342  VTNTLTFYVSSAVTILVNAVSSSR-----KLFL-----ESIGSLQGE 379

```

RESULT 5
T15816
hypothetical protein C48C5.1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_Revision 20-Sep-1999 #text_change 17-Mar-2000
C:Accession: T15816
R:Ravelli, A.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C48C5.
A:Reference number: T18410
A:Accession: T15816
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-378 <PAV>
A:Cross-references: EMBL:U39994, NID:G1055102, PID:G1055105, PIDN:AA837017.1, GSPDB:GN00
A:Experimental source: strain Bristol N2; clone C48C5
C:Genetics:
A:Gene: CESP:C48C5.1
A:Map position: X
A:Introns: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1
C:Superfamily: adenosine receptor A1

[illegible]

Db 320 QYMYFISGLEYLATTIINPIAYNLASSRFR-RAFKDIL 356

RESULT 6
157940

somatostatin receptor 5 - rat
 NAlternate names: somatostatin release-inhibiting factor subtype 28 receptor
 C1Species: Rattus norvegicus (Norway rat)
 C1date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 24-Nov-1996
 C1accession: I57400.1, I57400.2, 530201

R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C
Mol. Pharmacol. 42: 939-946, 1992

A>Title: Molecular cloning and expression of a pituitary somatostatin receptor with putative A1, A2 and A3 subtypes
A1,Reference number: I57940; MUID:93125499; PMID:1362243

A;Accession: 15/940
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA

A;Residues: 1-363 <OCAL>
A/Cross-references: GB:L04535; NID:g409238; PIDN:AAAI7029.1; PID:g409239
B/Occurrence: M: Total: 1 T: 1 M: 1

A1>Title: Molecular cloning and expression of a pituitary somatostatin receptor with pre

A:Reference number: 157949; MWID:94088493; PMID:9264565
A:Accession: 157949
Status: preliminary; translated from GR/EMBL/DBAT

A;Molecule type: mRNA
A;Residues: 341-363 <OCA2>

A/cross-references: GR:5613/0; MID:9455594//; PIDN:AAB29371.1; PID:94555948
A/Experimental source: pituitary
R/Penetta, R.; Greenwood, M.; Patel Y.C.

submitted to the EMBL Data Library, August 1993
A:Description: Correction of the nucleotide and amino acid sequence of the rat somatostatin
Reference number: S19244

A1:Accession: S39244
A1:Molecule type: mRNA

A:residues: 309-363 <PEN>
A:cross-references: EMBL:X74828; NID:g4333911; PIDN:CAA52825.1; PID:g4333912
C:Genetics:

A;Gene: SSTR5
C;Superfamily: vertebrate rhodopsin

Query Match	16.6%	Score 358;	DB 2;	Length 363;
Best Local Similarity	27.3%;	Pred. No. 1.9e-21;		

Matches	110;	Conservative	60;	Mismatches	149;	Indels	84;	Gaps	
5	WNGSDGPEGARBPMPAL	PCCDRRRCSPFPIG	---	IVPTAVCT	CT	FVVGVS	GNMTVM	61	

Db 12 WNSAASGNNH--WSLVG-----SASPMGARAVLVPEVLLVC--TVELSGNTLVIV 60

QY 62 LIGRYDMRTTNLYLGSMAVSDLLILGLPFDLYRLMRSRPWFVGPPLLCRLSLVYEGEC 121

Db 61 VLRHAKKKTVTNNVYILNLAVADVLEMLGLPFLATQNAVSVSYNEFGSFLCRLVMTLDGIN 120

```

121 QFTSIFCLMWSVDRIYLAVHPLRSRWRPRPVAKASAAVWYFSLMSLPFLVFADVOE 180C

```

QY 182 DPGISVVPGLNGTARIASSPLASSPELWLSRAPPPSPSGPETAEAALFSRECRPSPAQ 241

Db	181	G-----W-----	-----PVPVG	194
		-----WGTCNLS-----		

Db 195 LMGAFITYTTSVLGFEGLPLVICTCYLLIVKVKAAAGMRVGSSR-----RRRSEP 2444

QY 295 QTRKLVVLAFFICWLPFHVGRIIYN-TEDSRMIFYSQYENIVALQLPYLSASINPI 353

Db 245 KATRMVVVVVLVFVGCMPLFFIIVNIIVNLAFITLPEEPTSGAGLYFFVVLS - -YANSCANPL 302

303 LYGLFLSDNFRQSFRRKVLCLR---RGYGMEDADAIIEPRDDSG 341

RESULT 7

A46226

somatostatin receptor 3 - human

C/Species: Homo sapiens (man)

C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999

C/Accession: A46226; S32501

R/Yamada, Y.; Reisine, T.; Law, S.F.; Ihara, Y.; Kubota, A.; Kagimoto, S.; Seino, M.; Seino, M.; Endocrinol. 6, 2136-2142, 1992

A/Title: Somatostatin receptors, an expanding gene family: cloning and functional charac

A/Reference number: A46226; MUID:93149123; PMID:1337145

A/Accession: A46226

A/Molecule type: DNA

A/Residues: 1-418 <YAM>

A/Cross-references: GB:M66738; NID:9338498; PIDN:AAA60592.1; PID:9338499

A/Note: Sequence extracted from NCBI backbone (NCBIN:123685, NCBIPI:123690)

R/Corneff, J.D.; Demchyshyn, L.L.; Seeman, P.; van Tol, H.H.M.; Strikant, C.B.; Kent, G.; FEBS Lett. 321, 279-284, 1993

A/Title: A human somatostatin receptor (SSTR3), located on chromosome 22, displays pref

A/Reference number: S32501; MUID:9328970; PMID:8097479

A/Accession: S32501

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-418 <COR>

C/Genetics:

A:Gene: GDB:SSTR3

A/Cross-references: GDB:134187; OMIM:182453

A/Map position: 22q13.1-22q13.1

A/Introns: #status absent

C/Superfamily: vertebrate rhodopsin

F:44-70/Domain: transmembrane #status predicted <TM1>

F:81-106/Domain: transmembrane #status predicted <TM2>

F:118-139/Domain: transmembrane #status predicted <TM3>

F:159-181/Domain: transmembrane #status predicted <TM4>

F:203-233/Domain: transmembrane #status predicted <TM5>

F:255-282/Domain: transmembrane #status predicted <TM6>

F:289-316/Domain: transmembrane #status predicted <TM7>

F:17-30/Binding site: carbohydrate (asn) (covalent) #status predicted

F:116-191/Disulfide bonds: #status predicted

F:151-251,317,332/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status

F:256/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predic

F:412/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 16.5%; Score 356.5; DB 2; Length 418;

Best Local Similarity 26.9%; Pred. No. 3e-21;

Matches 100; Conservative 63; Mismatches 150; Indels 59; Gaps 9;

QY 11 PEGAREPPWALPPCDERRCSPPFL-----GALVPTAVCLFVVGSGNVVTMLIGR 65
 DB 15 PENA-SSAMPDPATLGNVSAQSPAGLAVSGVLLPVLVVC--VGLGSLVITVYLR 71
 QY 66 YRDMRTTNLYLGSMAVSDLLILGLPPDLVLRMSRPWFGLPCLRLSLVYGGCTYAT 125
 DB 72 HTASPSVTNYIINLADELFLMGLPF-LAAQNALSYWPGSLMCRVMAVDGINFTS 130
 QY 126 LHMHTALSVRYLAICRPLRAVLTTRRRVALLAVLMAVALLSAGPPLFLVGVQDDGI 185
 DB 131 IFCLTVMSVRYLAIVHPTSRARWTAIPVARTSAAVVAVVLPVAVSGV----- 184
 QY 186 SVVPGINGTARIASSPLASPLMLSRAPPSPPSGPETAALFSECRSPPAQLGAL 245
 DB 185 -----FRGMSHCIMQ-----WPEPAAWRA 204
 QY 246 RWMVWTTAVFPLPCLSLILYGLIGELWSSRRPLRGPASGRGRQTKRVLLVVL 305
 DB 205 GFIIVTAALFEFGPLVLCCLYLLIYVVRSAGRVMAFSCQRRRSRRVTRMVAVVA 264
 QY 306 AFIIICMLPHVGRITTYINTE-DSMMYHSQYFNVALQFLVLSASINILNLISKKYRA 364
 DB 265 LFLVLCMPFVYLVNVVCPLEBPAPFGILVVALP--YANSCANILYGFISYRK- 321

QY 365 AAPKLLARKSR 376
 DB 322 QGFRVLLRPSR 333

RESULT 8

I57955

somatostatin receptor - human

C/Species: Homo sapiens (man)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999

C/Accession: I57955

R/Panetta, R.; Greenwood, M.T.; Warynska, A.; Demchyshyn, L.L.; Day, R.; Niznik, H.B.

Mol. Pharmacol. 45, 417-427, 1994

A/Title: Molecular cloning, functional characterization, and chromosomal localization c

A/Reference number: I57955; MUID:94195267; PMID:7908405

A/Accession: I57955

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-363 <RES>

A/Cross-references: GB:U14865; NID:9431094; PIDN:AAA20828.1; PID:9431095

C/Genetics:

A:Gene: GDB:SST

A/Cross-references: GDB:119604; OMIM:182450

A/Map position: 3q28-3q28

C/Superfamily: vertebrate rhodopsin

Query Match 15.9%; Score 343; DB 2; Length 363;

Best Local Similarity 27.7%; Pred. No. 3.1e-20;

Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;

QY 14 AREPPWALPP-----CDERR-CSPPF-LGA--LVPTAVCLFVVGSGNVVTML 62
 DB 7 ASTPSWNASPGAAAGGDNRTLVPAPAGRAVLVPLVLLVC--AAGLGNTLVTV 64
 QY 63 IGRVDMRTTNLYLGSMAVSDLLILGLPPDLVLRMSRPWFGLPCLRLSLVYGGCT 122
 DB 65 VLPRAKKTVNTNYILNLAADVLMGLPF-LAQNAAAFPPFGPVLRCRLVMTLDVNG 123
 QY 123 YATLHMTALSVRYLAICRPLRAVLTTRRRVALLAVLMAVALLSAGPPLFLVGVQDD 182
 DB 124 FTSVVCULVMSVDRLAVLVHPLSSARWRPRVAKLASAAVWLSCMSPLLVADVQ- 182
 QY 183 PGISVPGINGTARIASSPLASPLMLSRAPPSPPSGPETAALFSECRSPPAQL 242
 DB 183 -----GGTVNAS-----W-----PPVGL 196
 QY 243 GALRMVWTTAVFPLPCLSLILYGLIGELWSSRRPLRGPASG-----RERGRQT 296
 DB 197 WGAFTITVAVLGFAPLVLVLCCLLVYKV-----RAACVRGCVRRRERKV 246
 QY 297 KRVLLVVALFIIICMLPHVGRITTYINTE-DSMMYHSQYFNVALQFLVLSASIN 351
 DB 247 TRMVLVVLVFAAGCWLPPFTVNTVNLVAVLPOEPASAGLYF--FVVI--LSYANSCAN 300
 QY 352 PLVNLISKKYRAAFAKLLARK 374
 DB 301 PVLVGLSDNFRQSFQKVLCLRK 323

RESULT 9

JN0763

somatostatin receptor 5 - human

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000

C/Accession: JN0763

R/Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.;

Biochem. Biophys. Res. Commun. 195, 844-852, 1993

A/Title: Cloning, functional expression and pharmacological characterization of a four

A/Reference number: JN0762; MUID:93384611; PMID:8373420

A/Accession: JN0763

A/Molecule type: DNA

A/Residues: 1-364 <YAM>

A/Cross-references: DDBJ:D16827; NID:9487683; PIDN:BA04107.1; PID:9487684

C:Comment: This protein is a member of somatostatin receptor family.

C:Genetics:

A:Gene: GDB:SSTR5

A:Cross-references: GDB:138452; OMIM:182455

A:Map position: 16p13.3-16p13.3

A:Introns: #status absent

C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; phosphoprotein; chlo

F:40-66/Domain: transmembrane #status predicted <TM1>

F:77-102/Domain: transmembrane #status predicted <TM2>

F:114-135/Domain: transmembrane #status predicted <TM3>

F:155-177/Domain: transmembrane #status predicted <TM4>

F:196-228/Domain: transmembrane #status predicted <TM5>

F:246-273/Domain: transmembrane #status predicted <TM6>

F:280-307/Domain: transmembrane #status predicted <TM7>

F:13-26,187/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:112-186/Disulfide bonds: #status predicted

F:242,325/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pr

F:247/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predic

F:320/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 15.9%; Score 343; DB 2; Length 364;

Best Local Similarity 27.7%; Pred. No. 3.2e-20;

Matches 106; Conservative 58; Mismatches 131; Indels 88; Gaps 14;

QY 14 AREPPPALP-----CDERR-CSPFP-LGA--LVPTAVCLCLFVVGSGVNTVNL 62

DB 7 ASPTSMWASSPGASGGGDNRTLVGPASAGARALVLPVLYLVC--AAGLGNTLVIV 64

QY 63 IGRYDRMTTNLVLGSAVSDLLILGLPFDLYRLMSRPVNPRLCRLSLVYGECT 122

DB 65 VLRPAKKTKVTNITLNLAVDLVLMGLP-LATONASFPFGPVLCRLVMTLDGVNQ 123

QY 123 YATLHMTALSVRYLAICRPLARVLTTRRRVRLALVLAVALISGPFPLVGEOD 182

DB 124 FTSVFCILTVMSVDYLLAVHPLSSARWRPRVAKLASAAWLSLMSIPLVFDVOC- 182

QY 183 PGISVVDGLNGTARIASSPPLWSRAPPSPGPETAALALFSRECRPSAQL 242

DB 183 -----GGTCNAS-----W-----DEPVGL 196

QY 243 GALRVMLVTTAVFELPPLCLSLYGLIGRELMSRRPLRGPASG-----RENGHQT 296

DB 197 WGAFTITVAVLGFAPLVLVCLCLLVIVK-----RAAGVRVGCYRRRSERV 246

QY 297 KRVLLVVLVAFIICMLPFHVGRIIYI-----NTEDSRMYFSQYFNIVALQLFYLASIN 351

DB 247 TRVVLVVLVAFAGCMLEPFTVNTVLAVALQEPASAGLVF---FVVT---LSYANSCAN 300

QY 352 PILYNLSKRYRAAFKLLARK 374

DB 301 PVLVGLSDNFRQSFQVCLARK 323

RESULT 10
A:Accession: A55259
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <XIE>
A:Cross-references: GB:U04092; NID:G476106; PIDN:AAA67171.1; PID:G476107
C:Superfamily: vertebrate rhodopsin
C:Keywords: transmembrane protein

Query Match 15.7%; Score 338; DB 2; Length 352;

Best Local Similarity 27.8%; Pred. No. 7.7e-20;

Matches 108; Conservative 52; Mismatches 109; Indels 106; Gaps 13;

QY 41 VTAVCCLFV--VGSGNVTVMLIGRYDMRTTNLYGSAVSDLLIL--GLPFDLY 96

DB 23 VVSFVLVLLVCTLTGIGNMVLVLTSDMTPTNCYLSLALADLVLAAGLPVSD 82

QY 97 RLMRSPPWFVGLPCLSLYVGECCYATLL-----HMTALSVRYLAICRPLARVL 149

DB 83 SL--VGHMTVGRAGCL-----GLTYFOYLGINVSSFSLIAFVRYLAICRPLARQTV 133

QY 150 VTRRRVRLALVLAVALISGPFPLVGV-----EODPGISVVDGLNGTARIASSPLA 203

DB 134 CVAAAKRIIAIGIKWTSYCLMLPFLVDLVNRDNRLECGYVPPGL----- 181

QY 204 SSPPLWSRAPPSPPSGPETAALALFSRECRPSAQLGALRVMLVTTAVFELPCL 263

Query Match 15.9%; Score 342.5; DB 2; Length 380;

Best Local Similarity 25.1%; Pred. No. 3.6e-20;

Matches 102; Conservative 74; Mismatches 147; Indels 83; Gaps 12;

QY 2 GSPW-----NGSDPGGAREPP--WPALEPCDERCRSPPLGALVTVAVCLCF 49

DB 26 GSAMLPGMABPDGAGAGQDEQLLEPAHISPAIP-----VIITVYSVVF 70

QY 50 VVSGNVTVMLIGRYDRMTTNLYGSAVSDLLILGLPFDLYRLMSRPVNPRL 109

DB 71 VVGLVGNLSLVMPFIITRTKTKATNTIYPLALADALVTTMPQ--STVYLSMSPFEDV 129

QY 110 LCRSLYVGECCYATLLHMTALSVRYLAICRPLARVLTTRRRVRLALVLAVALIS 169

DB 130 LCKIVISIDVYNNFTSIFLTLMMSVDYLLAVHPLSSARWRPRVAKLASAAWLSLMSIPLVFDVOC- 182

QY 170 AGFPLVGVGEQPGISVVDGLNGTARIASSPPLWSRAPPSPGPETAALAA 229

DB 190 GISAILLGTRKREDVDIIE-----CSLQFDDDDYSWD----- 223

QY 230 LFSRECRPSAQLGALRVMLVTTAVFELPPLCLSLYGLIGRELMSRRPLRGPASGR 289

DB 224 LFWKIC-----VVPAPVIVLIIIVCTYTLMIIRL-KSVLLSG--SREK 265

QY 290 ERGHRQTKRVLLVLAFLIICMLPFHVGRIIYI--NTEDSRMYFSQYFNIVALQLFYL 347

DB 266 DRNLRIITRLVAVVAVFIICMTPIHIFLVEALGSTSHSTAALSSYF--CIALGYTN 322

QY 348 ASINPLIYNLSKRYRAA-----AFLLAKSRPPGFRSRDTA 387

DB 323 SSLLNPILVAFLENDFKRCFDFCPFKMREROSTSRVNTVDDPA 368

RESULT 11

JE0296
Thyrotropin releasing hormone receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000

C:Accession: JE0296

R:Itahara, H.; Nakamura, T.; Itob, J.; Iwasa, H.; Kanatani, A.; Borkowski, J.; Ihara,

Biochem. Biophys. Res. Commun. 250, 68-71, 1998

A:Title: Cloning and characterization of a new subtype of thyrotropin-releasing hormone

A:Reference number: JE0296; MUID:98407892; PMID:9735333

A:Accession: JE0296

A:Molecule type: mRNA

A:Residues: 1-352 <ITTA>

A:Cross-references: DDBJ:AB015645; NID:G3660553; PIDN:BAA33437.1; PID:G3660554

C:Superfamily: adenosine receptor A1

F:26-48/Domain: transmembrane #status predicted <TM1>

F:58-80/Domain: transmembrane #status predicted <TM2>

F:97-118/Domain: transmembrane #status predicted <TM3>

F:142-165/Domain: transmembrane #status predicted <TM4>

F:188-209/Domain: transmembrane #status predicted <TM5>

F:252-273/Domain: transmembrane #status predicted <TM6>

F:282-304/Domain: transmembrane #status predicted <TM7>

Query Match 15.7%; Score 338; DB 2; Length 352;

Best Local Similarity 27.8%; Pred. No. 7.7e-20;

Matches 108; Conservative 52; Mismatches 109; Indels 106; Gaps 13;

QY 41 VTAVCCLFV--VGSGNVTVMLIGRYDMRTTNLYGSAVSDLLIL--GLPFDLY 96

DB 23 VVSFVLVLLVCTLTGIGNMVLVLTSDMTPTNCYLSLALADLVLAAGLPVSD 82

QY 97 RLMRSPPWFVGLPCLSLYVGECCYATLL-----HMTALSVRYLAICRPLARVL 149

DB 83 SL--VGHMTVGRAGCL-----GLTYFOYLGINVSSFSLIAFVRYLAICRPLARQTV 133

QY 150 VTRRRVRLALVLAVALISGPFPLVGV-----EODPGISVVDGLNGTARIASSPLA 203

DB 134 CVAAAKRIIAIGIKWTSYCLMLPFLVDLVNRDNRLECGYVPPGL----- 181

QY 204 SSPPLWSRAPPSPPSGPETAALALFSRECRPSAQLGALRVMLVTTAVFELPCL 263

Db 182 ----- YIPYILDFAVFIQGLLVT 201
Oy 264 SILYGLIGR-----ELWSSRRPLRG--PAASGR---ERGRQTRKVLVVLAFT 308
Db 202 LVLYGLIGRLPQSPLSQAMQKEROQPHQGEAPAGNSRAKSSRQATRMALVAVLLFA 261
Oy 309 ICWLPHVHGRIIYINTEDSRMMYFSQVFNIVALQLF-----YLSASINFLNLYLSKKYR 363
Db 262 VLMTFPYRT--LVTLNS-----FVAQFPLDPLVLLFCRTCVYTNSAVNNEVVSLSQKFR 313
Oy 364 AAAFKLKWLAKSRPR 378
Db 314 AAFLKLCWCRAPAGPQ 328
RESULT 12
S36143
kappa opioid receptor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Dec-1993 #sequence revision 19-0ct-1995 #text_change 20-Jun-2000
C/Accession: S36143; S38825; S36102; S39015; A48789
R/Nishi, M.; Takeshima, H.; Fukuda, K.; Kato, S.; Mori, K.
FEBS Lett. 330, 77-80, 1993
A>Title: cDNA cloning and pharmacological characterization of an opioid receptor with hi
A/Reference number: S36143; MUID:93380575; PMID:9396539
A/Accession: S36143
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-380 <NIS>
A/Cross-references: GB:D16534; NID:g409390; PIDN:BAA03971.1; PID:g415310
R/Chen, Y.; Mestek, A.; Liu, J.; Yu, L.
Biochem. J. 295, 625-628, 1993
A>Title: Molecular cloning of a rat kappa opioid receptor reveals sequence similarities
A/Reference number: S38825; MUID:94059008; PMID:8240267
A/Accession: S38825
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-380 <CHE>
A/Cross-references: GB:D22001; NID:g409236; PIDN:AAA41495.1; PID:g409237
R/Minami, M.; Toya, T.; Katso, Y.; Maekawa, K.; Nakamura, S.; Onogi, T.; Kaneko, S.; Sai
FEBS Lett. 329, 291-295, 1993
A>Title: Cloning and expression of a cDNA for the rat kappa-opioid receptor.
A/Reference number: S36102; MUID:93374033; PMID:8103466
A/Accession: S36102
A/Molecule type: mRNA
A/Residues: 1-41, 'L', 43-380 <MIN>
A/Cross-references: GB:D16829; NID:g404115; PIDN:BAA04109.1; PID:g404116
R/Li, S.; Zhu, J.; Chen, C.; Chen, Y.W.; Deriel, J.K.; Ashby, B.; Liu-Chen, L.Y.
Biochem. J. 295, 629-633, 1993
A>Title: Molecular cloning and expression of a rat kappa opioid receptor.
A/Reference number: S39015; MUID:94059009; PMID:8240268
A/Accession: S39015
A/Molecule type: mRNA
A/Residues: 1-344, 'Y', 346-380 <NIS>
A/Cross-references: GB:L22536; NID:g425188; PIDN:AAA41496.1; PID:g425189
R/Meng, F.; Xie, G.
Proc. Natl. Acad. Sci. U.S.A. 90, 9954-9958, 1993
A>Title: Cloning and pharmacological characterization of a rat kappa opioid receptor.
A/Reference number: A48789; MUID:94052210; PMID:8234341
A/Accession: A48789
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-380 <RES>
A/Cross-references: EMBL:U00442; NID:g403486; PIDN:AAA18261.1; PID:g403487
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein
Query Match 15.7%; Score 338; DB 2; Length 380;
Best Local Similarity 25.5%; Pred. No. 8, 4e-20;
Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;
6 NGSGCPGAGARPP---WPLPDCDERRCSPFGLGALVPAVAVCLFLFVGVSGNVVTML 62

Db 39 NGSVSEDOQLPEAHISPAIP-----VITTAIVSYVFPVGLGNSLWFEV 83
QY 63 IGRYDRAFTTNLYIGSMAVSDLLLLGLPFLYRLMTSRPVPFGPLCRSLAYGEGCT 122
Db 84 IIRYKMTATNTIYIFNLALADALVTTMPFQ-SAVYLMNSMPFGDLCKIYISIDYNN 142
QY 123 YATLLHMTALSYERLTAICRPLARVLVTRRRVRLALINLWALLSAGPFLFLVGEOD 182
Db 143 FTSITLLTMASVDRIANCHPVKALDFRPLKAKIINICIMWLLASVGSALVLGGTKVR 202
QY 183 PGISVYPGINGTARIASSEPLASSPPLWLISRAPSPSPSGSEFTAEAAALFSRECRPSPAOL 242
Db 203 EDVDVIE-----CSLQGFDPDEXSMWD-----LFMKIC----- 229
QY 243 GALRYMLWTTAYVFLPLFLCLSLYGLIGRELMSRRPLRGPAASGRGHHQTKRVLV 302
Db 230 -----VFVEAFVIPVLIIIVCYTLMTLRL-KSVRLISG--SREKDNLRITLVLV 278
QY 303 VVLALFLICLPHVGRHIIYV-NTEDSRMYPPOXYFENIYALOLFYLASINPLIYNLISK 360
Db 279 VVAVFILICMTPIPHIFLVALGSTSHSTVLSYYR---CINALGYNSSLNPLVLAFLDE 335
QY 361 KYRAA-----AFKLLARKSRPRGFHRSRDTAGEVAG--DTGG 396
Db 336 NKRCRFRDPCFPIKXMERQST---NRVNTYQDPASMRDVG 375
RESULT 13
JC2338
kappa opioid receptor 1 - human
C.Species: Homo sapiens (man)
C.Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #ext_change 19-May-2000
C.Accession: JC2338; A55354; I57005; G01546
R.Manson, E.; Bare, L.; Yang, D.
Biochem. Biophys. Res. Commun. 202, 1431-1437, 1994
A.Title: Isolation of a human kappa opioid receptor cDNA from placenta.
A.Reference number: JC2338; MUID:9438360; PMID:8060324
A.Accession: JC2338
A.Molecule type: mRNA
A.Residues: 1-380 <MAN>
A.Cross-references: GI:U1053; NID:G532059; PIDN:AAA20985.1; PID:G532060
A.Experimental source: placenta
R.Wang, J.B.; Johnson, P.S.; Wu, J.M.; Wang, W.F.; Uhl, G.R.
J. Biol. Chem. 269, 25966-25969, 1994
A.Title: Human kappa opiate receptor second extracellular loop elevates dynorphin's affinity.
A.Reference number: A55354; MUID:95014415; PMID:7929306
A.Accession: A55354
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 136-279 <MAN>
A.Cross-references: GB:U136130; NID:G598184; PIDN:AAA63646.1; PID:G598185
R.Zhu, J.; Chen, C.; Xue, J.
Life Sci. 56, 201-207, 1995
A.Title: Cloning of a human kappa opioid receptor from the brain.
A.Reference number: I57005
A.Accession: I57005
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1/'D',3-380 <ZHU>
A.Cross-references: GB:U137362; NID:G722617; PIDN:AAA63906.1; PID:G722618
R.Grandy, D.K.
submitted to the EMBL Data Library, November 1994
A.Reference number: G07718
A.Accession: G01546
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 132-203 <GRA>
A.Cross-references: EMBL:U16860; NID:G595932; PIDN:AAA56758.1; PID:G595933
C.Comment: This receptor preferentially binds dynorphins.
C.Genetics:
A.Gene: GDB:OPRK1; KOR
A.Cross-references: GDB:1132651; OMIM:165196

A:Map position: 8q11.2-8q11.2
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; receptor; transmembrane protein
 F:60-85/Domain: transmembrane #status predicted <TM2>
 F:95-114/Domain: transmembrane #status predicted <TM2>
 F:133-154/Domain: transmembrane #status predicted <TM3>
 F:177-199/Domain: transmembrane #status predicted <TM4>
 F:228-251/Domain: transmembrane #status predicted <TM5>
 F:275-296/Domain: transmembrane #status predicted <TM6>
 F:311-333/Domain: transmembrane #status predicted <TM7>

Query Match 15.5%; Score 334; DB 2; Length 380;
 Best Local Similarity 25.2%; Pred. No. 1.8e-19;
 Matches 99; Conservative 74; Mismatches 146; Indels 74; Gaps 11;

QY 6 NGSDGPEGAREPP---WPALPDCERRCSPFPLGALVPTAVCLCFVVGSGNVTVML 62
 Db 39 NGASGSDAQLEPAHISPAIP-----VITAVSVFVVGSLVMSLVFV 83

QY 63 IGRYDMRTTNLYGSAVSDLLILGLPDLYLMSRPWFVGLLCLSLYVGEQCT 122
 Db 84 IIRYKMKATATYIYFNLLADALVTTMPFQ-STVYLMNSWPFQVLCXIVISIDVYM 142

QY 123 YATLLHMTALSVRYLATCRPLRARVLTTRRRALAVLMVALLSAGFLVGVED 182
 Db 143 FTSITTLTMSVDRYIAVCHPVKALDFRTPLKAKITNICWILSSVGSIAVLGGTKR 202

QY 183 PGISVPLGINTARIASPELASSPPLMLSRAPPSPSGPETAAMALFSRECRPSPAQL 242
 Db 203 EDVPIE-----CSIQFPDDDYSMWD-----LPMKIC----- 229

QY 243 GALKRMLWTTTYFPLPCLSLYLGLIRELMSRRPLRGPAASRGRRGHRQTKVLLV 302
 Db 230 -----VFIFPAFVPLVLIIVCYLMLRL-KSVRLLSG--SRKEDNLRITRLVLV 278

QY 303 VLAFLICMLPEHVGRIYI-NTEDSRMVFQYFNIVALOLFYSASINPLVNLISK 360
 Db 279 VVAIVFVCMTPHIFILVEALGSTSHSTALSSYF--CIALGYTSSLNPLVAFLE 335

QY 361 KYRAA-----AFKLLLRKSRPRGFRSDTA 387
 Db 336 NFKRCFRDPCFPLKWMERQSTSRVNTVQDPA 368

RESULT 14
 JN0605
 somatostatin receptor 4 - human
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 20-Jun-2000
 C:Accession: JN0605; JN0762; A47457
 R:Ku, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.
 Biochem. Biophys. Res. Commun. 193, 648-652, 1993
 A:Title: Molecular cloning and sequencing of a human somatostatin receptor, hsstr4.
 A:Reference number: JN0605; MUID:93290656; PMID:8515564
 A:Accession: JN0605
 A:Molecule type: DNA
 A:Residues: 1-388 <XUY>
 A:Cross-references: GB:144856; NID:9292499; PIDN:AAA6623.1; PID:9292500
 R:Yamada, Y.; Kagitcho, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I
 Biochem. Biophys. Res. Commun. 195, 844-852, 1993
 A:Title: Cloning, functional expression and pharmacological characterization of a fourth
 A:Reference number: JN0762; MUID:93384611; PMID:8373420
 A:Accession: JN0762
 A:Molecule type: DNA
 A:Residues: 1-388 <YMA>
 A:Cross-references: GB:16826; NID:9693907; PIDN:BA04106.1; PID:9693908
 R:Rohrer, L.; Raulf, F.; Bruns, C.; Buettnr, R.; Hofstaedter, F.; Schulte, R.
 Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
 A:Title: Cloning and characterization of a fourth human somatostatin receptor.
 A:Reference number: A47457; MUID:93248256; PMID:8483934
 A:Accession: A47457
 A:Molecule type: DNA
 A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>

A:Cross-references: GB:107833; NID:9307429; PIDN:AAA60565.1; PID:9307430
 A>Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBI:P:130858)
 C:Comment: This protein mediates the diverse actions of the tetradecapeptide somatostatin
 C:Genetics:
 A:Gene: GDB:SSTR4
 A:Cross-references: GDB:202662; OMIM:182454
 A:Map position: 20p11.2-20p11.2
 A:Introns: #status absent
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; ph
 F:47-73/Domain: transmembrane #status predicted <TM1>
 F:84-109/Domain: transmembrane #status predicted <TM2>
 F:121-142/Domain: transmembrane #status predicted <TM3>
 F:162-184/Domain: transmembrane #status predicted <TM4>
 F:208-238/Domain: transmembrane #status predicted <TM5>
 F:257-284/Domain: transmembrane #status predicted <TM6>
 F:291-314/Domain: transmembrane #status predicted <TM7>
 F:324/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:119-198/Disulfide bonds: #status predicted
 F:161,253/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status
 F:327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 15.5%; Score 334; DB 2; Length 388;
 Best Local Similarity 26.9%; Pred. No. 1.8e-19;
 Matches 105; Conservative 65; Mismatches 128; Indels 92; Gaps 15;

QY 1 MGSPW---NGSDGPEGAREPPWPALPDCERRCSPFPLGALVPTAVCLCFVVGSGN 56
 Db 15 LGTAMPANAASSAPABAE---AVAGBDAR-----AAGVIAICITVALCVLGVGN 65

QY 57 VVTVMILGRYDMRTTNLYGSAVSDLLILGLPDLYLMSRPWFVGLLCLSLY 116
 Db 66 ALVIFVILRYAKMKATATYIYFNLLADALVTTMPFQ-STVYLMNSWPFQVLCXIVISIDVYM 124

QY 117 VGBGCTATLLHMTALSVRYLATCRPLRARVLTTRRR--VRALIAV-LMAVALLSAGFL 173
 Db 125 VDGLNMTSFVCLVLSVDRVAVVHPLRA--ATYRPSVAKLINIGWLASLVLTPFI 181

QY 174 LFLVGVEDPGISVPLGINTARIASPELASSPPLMLSRAPPSPSGPETAAMALFSR 233
 Db 182 AIFADTR-----PARGQA--VACNLQWPHAW-----SAVF-- 211

QY 234 ECRPSPAQLGALKRMLWTTTYFPLPCLSLYLGLIREL-----MSRRPLRGPA 285
 Db 212 -----VVYTFPLGLFPLVLAIGLCYLLIVGKRAVALRAGWQQR----- 251

QY 286 ASRGRRGHRQTKVLLVYVLAFLICMLPEHVGRI--IYINTEDSRMVFQYFNIVALOL 343
 Db 252 -----NSEKRTPLVLMVAVVFLCWPFFVVOQLNLVWTSLDATV-----NHSVLL 299

QY 344 FYLSASINPLVNLISKRYRAAFKLLAR 373
 Db 300 SVANSCANPLTYGFLSDNPFRRSRQVLCR 329

RESULT 15
 I49022
 kappa opioid receptor 3 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text_change 19-May-2000
 C:Accession: I49022
 R:Pam, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brooks, A.I.; D
 Mol. Pharmacol. 47, 1180-1188, 1995
 A:Title: Cloning and functional characterization through antisense mapping of a kappa 3
 A:Reference number: I49022; MUID:95327076; PMID:7603458
 A:Accession: I49022
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RES>
 A:Cross-references: EMBL:U09421; NID:9551484; PIDN:AAA8133.1; PID:9551485
 C:Superfamily: vertebrate rhodopsin

Query Match 15.5%; Score 333.5; DB 2; Length 367;

Best Local Similarity 27.1%; Pred. No. 1.9e-19;
Matches 101; Conservative 67; Mismatches 132; Indels 73; Gaps 12;

```

QY 34 ELGALVPTAVCLCLEVVGVSANVTMLIGRYRDMRTTNLYLSMAVSDLLILGLPF 93
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 44 FLGLKVTIVGLYLAIVCIGLLGNCLVMYVILRHTRKMTATNIIYFNIALADTLVLLTLP 103
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 94 ---DIYRLMRSPWVFGPILCRSLVVGEGCTYATLLHMTLSVERYLAICRPLRVLV 150
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 104 QGTDLILGF---WPGNALCKIVIAIDYNNMFTSTFTLTAMSVDRYVAICHPIRALDVR 159
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 151 TFRRYVALIYLMVAVALLSAGPFLPLVGVEQDPGISVPGINGTARIASSPLASSPPLWL 210
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 TSSKQAQVAVVAIMWALASV-----VGVPVAIMGSAQVEDEIEC-----L 198
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 SRAPPPSPSPGPETAEALAFSRECRPSPAQLGALRMVMTTAYFFLPFLCLSLYGLI 270
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 VEIPAPQDYWGCVFAICIFLFS-----FLIPVLIISVCYSIM 235
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 271 GRELWSSRRPLGPAASGRERGHROTGRVLLVYVLAFLICMLPFHVGRITTYINTEDSRMM 330
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 IKRL-RGVRLISG--SREKDRNLRIITRLVAVVAVFVGCWTEPVQV---FVLVQSLGVQ 268
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 331 YFSQYFNIVALQLF-----YLSASINPIILYNISSKYRAAFKILLARKSRPRGFHRS-- 383
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 269 PGSE--TAVAILRFCTALGVVNSCLNPITLYAFIDENFKACFRKCCA-----SALHREMQ 341
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 -RDTAGEVAGDTG 395
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 342 VSDRVRTVAKDVG 354
    ||| : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: March 16, 2003, 15:05:27
Job time : 20 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:04:24 ; Search time 14 Seconds

(without alignments)
1356.424 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156
Sequence: 1 MGSPMNGSDGPGAGARPPWP.....DTGGDTVGYTTSANTKTMG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	430	19.9	419	10	US-09-804-551B-26
2	369.5	17.1	362	9	US-09-992-331-13
3	358	16.6	363	9	US-09-992-331-14
4	356.5	16.5	418	9	US-09-992-331-17
5	355.5	16.5	370	9	US-09-966-782A-13
6	349.5	16.2	453	9	US-09-970-966-209
7	349.5	16.2	453	9	US-09-825-294-209
8	340	15.8	380	10	US-09-966-871-77
9	340	15.8	380	12	US-10-039-645-77
10	336.5	15.6	376	10	US-09-966-871-76
11	336.5	15.6	376	12	US-10-039-645-76
12	335.5	15.6	367	9	US-09-823-114-23
13	333.5	15.5	367	9	US-10-087-345A-15
14	333.5	15.5	428	9	US-09-992-331-16
15	331	15.4	372	9	US-09-966-782A-16
16	331	15.4	372	9	US-10-212-980-9
17	330.5	15.3	372	10	US-09-966-871-80
18	330.5	15.3	372	12	US-10-039-645-80
19	330.5	15.3	428	9	US-09-992-331-15

20	330	15.3	376	10	US-09-823-114-17	Sequence 17, Appl
21	330	15.3	394	10	US-09-823-114-22	Sequence 22, Appl
22	329.5	15.3	367	10	US-09-170-519-4	Sequence 4, Appl
23	329.5	15.3	370	10	US-09-823-114-19	Sequence 19, Appl
24	328	15.2	380	10	US-09-214-804-6	Sequence 6, Appl
25	326.5	15.1	371	9	US-09-966-782A-15	Sequence 15, Appl
26	326.5	15.1	371	9	US-10-212-980-8	Sequence 8, Appl
27	325.5	15.1	372	9	US-10-112-559A-4	Sequence 4, Appl
28	321.5	14.9	370	10	US-09-823-114-21	Sequence 21, Appl
29	320.5	14.9	368	9	US-10-212-980-2	Sequence 2, Appl
30	319.5	14.8	368	9	US-09-966-782A-14	Sequence 14, Appl
31	319.5	14.8	368	9	US-10-212-980-4	Sequence 4, Appl
32	319.5	14.8	427	9	US-10-007-132-4	Sequence 4, Appl
33	319	14.8	372	9	US-10-112-559A-2	Sequence 2, Appl
34	319	14.8	372	10	US-09-214-804-4	Sequence 4, Appl
35	319	14.8	380	9	US-10-166-359-2	Sequence 2, Appl
36	319	14.8	380	9	US-10-166-113-2	Sequence 2, Appl
37	319	14.8	380	9	US-10-166-357-2	Sequence 2, Appl
38	319	14.8	380	9	US-10-166-372-2	Sequence 2, Appl
39	319	14.8	380	10	US-09-149-045-2	Sequence 2, Appl
40	318.5	14.8	387	9	US-09-966-782A-17	Sequence 17, Appl
41	318.5	14.8	387	9	US-10-212-980-10	Sequence 10, Appl
42	317.5	14.7	370	9	US-10-007-132-2	Sequence 2, Appl
43	316	14.7	362	9	US-10-087-345A-17	Sequence 17, Appl
44	315	14.6	383	10	US-09-966-871-87	Sequence 87, Appl
45	315	14.6	383	12	US-10-039-645-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-09-804-551B-26
; Sequence 26, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Ia A 34 394
; CURRENT APPLICATION NUMBER: US/09/804,551B
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: DE 100 13 618.4
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-804-551B-26

Query Match 19.9%; Score 430; DB 10; Length 419;
Best Local Similarity 27.7%; Pred. No. 1.4e-27;
Matches 119; Conservative 61; Mismatches 119; Indels 130; Gaps 10;

QY	23	PCDDRRCSFPGALVPTAVCLFLVGVSGNVVTMLIGRYDMRTTNLYGSMNV	82
DB	12	PPRD-----PLATVPVTVVSLFITGVGNISTCIVIKORSMTATNYYLFLSLAI	64
QY	83	SDLLIIL-GHPDLYRMRSPWFGPLCLSLVGBGCTYATILMHTALSVERYLAI	141
DB	65	SDFLLLSGVQEVSYWSKYFVFGYICIGRLMETSANAATVLTITATFVERVIAIC	124
QY	142	RLPRLVLRVRRVRLAIYLVMAVALSAGPFLVGVGQDPGTSVVGNGTARIASSP	201
DB	125	HPFLGQMSKLSAIRITIVVMAIYTAIPQAQFGEHHSVGE-----	169
QY	202	LASSPPLMWSRAPPPSPSPETAEAALFSRRCRPSAPQAGLRVWL---WVTAYAF	257
DB	170	-----CCGIYRVIVKHSFQSTFTFF	190
QY	258	L-PFLCISLYGLIGRLWSSRRPLRGPAASGRER-----	291

Db 191 LAPMSIIIVLYLIGVHLRSTL-VEGPASVVAROOLKSVSDTLTYRGSGTAMSFNG 249
 QY 232 -----GHRQKRVLLVVLAFIICMLPFPHVGRITTY- 322
 Db 250 GSGAGTAGIAGGGAQLSSVVRGLNHVGTFRVRLMVAVVGFFLCMAFFHQRLLAIY 309
 QY 323 -----NTEDSRMNYFSQYFNITVALDLYLSASINPILYLSKRYRAAFKLLARKSR 376
 Db 310 APARGAKLRDQHEFVYT-VMTYVSGLVYLSTICINPLINIMSHKFR-EAFKAVLFGKRV 367
 QY 377 PRGFHRSRD 385
 Db 368 SKGSLSNRN 376

RESULT 2
 US-09-992-331-13
 ; Sequence 13, Application US/09992331
 ; Publication No. US20030022186A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FEDER, JOHN N.
 ; APPLICANT: MINTIER, GABE
 ; APPLICANT: RAMANATHAN, CHANDRA S.
 ; APPLICANT: HAWKEN, DONALD R.
 ; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,
 ; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
 ; TITLE OF INVENTION: CELLS
 ; FILE REFERENCE: D0048NP
 ; CURRENT APPLICATION NUMBER: US/09/992,331
 ; CURRENT FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: 60/308,540
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: 60/261,782
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 60/248,483
 ; PRIOR FILING DATE: 2000-11-14
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 13
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-992-331-13

Query Match 17.1%; Score 369.5; DB 9; Length 362;
 Best Local Similarity 27.8%; Pred. No. 1e-22;
 Matches 112; Conservative 58; Mismatches 148; Indels 85; Gaps 13;

QY 5 WNGSDGEGAREPPMPALPDCERRCSFPLGA---LVPTAVCLCLFVGVSGNVVTVM 61
 Db 12 WNSAASGSHN--WSLVDPVS-----PMGARAVLPVLYLVC--TVGSGNTLVLY 60
 QY 62 LIGRYDMRTTNNLYLGSMAVSDLLILGLPPDLYRLMRSPWVFGPLLCRLSLYVGGC 121
 Db 61 VLRRAKKTIVNVIILNLAADVLFMLGLPFLATQNAVSWPFSFCLRLVMTLDGIN 119
 QY 122 TYATLHMTALSVRRYLAICRPLRARVLTERRRVALIAVMAVALLSAGPFLFVGVG 181
 Db 120 QFTSIFCLMWSVDRYLAHVHPLRSARWRPRVAKLASAAVWFSLSMLPLLVADVOE 179
 QY 182 DPGISVVEGLNGTARIASSPLASPPPLWSRAPPPSPGSETAEMAALFSRECRSPAQ 241
 Db 180 G-----WGTCNLS-----W-----PEPVG 193
 QY 242 LGALRVMMLWTTAVFELPLCLSLIYGLI-----GRELMSRRPLRGPASGERGRH 294
 Db 194 LMGAAFTITYSVLGFGLVLCICLYLIVKKAAGMRVGSRR-----RRRSEP 243
 QY 295 QTKRVLLVVLAFIICMLPFPHVGRITTYIN-TEDSRMNYFSQYFNITVALDLYLSASINP 353
 Db 244 KTRMNVVVVVLVFGCMLPFPIVNIINLAFTLPEEPTSGALYFVVVLS--YANSCANPL 301
 QY 354 LYNLSKRYRAAFKLLARKSRPRGFHRSRDTAGVAGDTG 396

Db 302 LYGFSLDNFROSFRKALCLR-----RGYGVEDADAIIEPRDKSG 340

RESULT 3
 US-09-992-331-14
 ; Sequence 14, Application US/09992331
 ; Publication No. US20030022186A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FEDER, JOHN N.
 ; APPLICANT: MINTIER, GABE
 ; APPLICANT: RAMANATHAN, CHANDRA S.
 ; APPLICANT: HAWKEN, DONALD R.
 ; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,
 ; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
 ; TITLE OF INVENTION: CELLS
 ; FILE REFERENCE: D0048NP
 ; CURRENT APPLICATION NUMBER: US/09/992,331
 ; CURRENT FILING DATE: 2001-11-14
 ; PRIOR APPLICATION NUMBER: 60/308,540
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: 60/261,782
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 60/248,483
 ; PRIOR FILING DATE: 2000-11-14
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 363
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; US-09-992-331-14

Query Match 16.6%; Score 358; DB 9; Length 363;
 Best Local Similarity 27.3%; Pred. No. 8.8e-22;
 Matches 110; Conservative 60; Mismatches 149; Indels 84; Gaps 12;

QY 5 WNGSDGEGAREPPMPALPDCERRCSFPLGA---LVPTAVCLCLFVGVSGNVVTVM 61
 Db 12 WNSAASGSHN--WSLVG-----SASPMGARAVLPVLYLVC--TVGSGNTLVLY 60
 QY 62 LIGRYDMRTTNNLYLGSMAVSDLLILGLPPDLYRLMRSPWVFGPLLCRLSLYVGGC 121
 Db 61 VLRRAKKTIVNVIILNLAADVLFMLGLPFLATQNAVSWPFSFCLRLVMTLDGIN 120
 QY 122 TYATLHMTALSVRRYLAICRPLRARVLTERRRVALIAVMAVALLSAGPFLFVGVG 181
 Db 121 QFTSIFCLMWSVDRYLAHVHPLRSARWRPRVAKLASAAVWFSLSMLPLLVADVOE 180
 QY 182 DPGISVVEGLNGTARIASSPLASPPPLWSRAPPPSPGSETAEMAALFSRECRSPAQ 241
 Db 181 G-----WGTCNLS-----W-----PEPVG 194
 QY 242 LGALRVMMLWTTAVFELPLCLSLIYGLI-----GRELMSRRPLRGPASGERGRH 294
 Db 195 LMGAAFTITYSVLGFGLVLCICLYLIVKKAAGMRVGSRR-----RRRSEP 244
 QY 295 QTKRVLLVVLAFIICMLPFPHVGRITTYIN-TEDSRMNYFSQYFNITVALDLYLSASINP 353
 Db 245 KTRMNVVVVVLVFGCMLPFPIVNIINLAFTLPEEPTSGALYFVVVLS--YANSCANPL 302
 QY 354 LYNLSKRYRAAFKLLARKSRPRGFHRSRDTAGVAGDTG 396
 Db 303 LYGFSLDNFROSFRKALCLR-----RGYGVEDADAIIEPRDKSG 341

RESULT 4
 US-09-992-331-17
 ; Sequence 17, Application US/09992331
 ; Publication No. US20030022186A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FEDER, JOHN N.
 ; APPLICANT: MINTIER, GABE

```

; APPLICANT: RAMANATHAN, CHANDRA S.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY18,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: D0048ND
; CURRENT APPLICATION NUMBER: US/09/992,331
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/308,540
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/261,782
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/248,483
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-331-17

```

Query Match 16.5%; Score 356.5; DB 9; Length 418;

Best Local Similarity 26.9%; Pred. No. 1.4e-21; Matches 100; Conservative 63; Mismatches 150; Indels 59; Gaps 9;

```

QY 11 PEGAREPWPALPCDERRCSPPL-----GALVPTAVCLCFVVGSGNVVTMLIGR 65
DB 15 PENASAMPDPAIGNVSAGSPSAGLAVSGVLIPLVIVVC--VYGLGNSLVIYVLR 71
QY 66 YRDMRTTNLYIGMAVSDLLILGLPDLVLMRSRPVFGGLCRSLSYVEGCTYAT 125
DB 72 HTASBSVTNVYILNLADLDELFLGLP--LAAQNALSYWPFSGMLVMAVDGINQFTS 130
QY 126 LHMHTALSVERYLACRPLARAVLTVRRVRALIAVMAVALLSAGPFLVGVSDPGI 185
DB 131 IFCLTVMSVDRVLAVVHTRTSARWRTAPARIVSAVAWVAVVLPVVVFSGV----- 184
QY 186 SVFGLNGTARIASSPLASSPPLWLSRAPSPSPGSETAEAAALFSRECRSPPAQLGAL 245
DB 185 -----PRGMSTCHNQ-----WPEPALAMBA 244
QY 246 RVMLWTTAYFELPFLCLSLYGLIGREIMSSRPLRGASGREGHQTKRVLLVVL 305
DB 205 GFTITTAALGFEPFLVLCYLLIVKVSAGRRVWASCCORRRSERRVTMVAVA 264
QY 306 AFITCWLPHVGRITTYNTE--DSRMVFSQYFNIVALQLFYLSASINPILYMLISKYRA 364
DB 265 LFLVCMWPFYVLTINIVVCPLEBEPAFGLYFLVVALP--YANSCANPILYGFSLYRFK- 321
QY 365 AAFKLLARKSR 376
DB 322 QGFRVLLRPSR 333

```

RESULT 5

```

US-09-966-782A-13
; Sequence 13, Application US/09966782A
; Publication No. US2003002183A1
; GENERAL INFORMATION:
; APPLICANT: Battaglin, P.
; APPLICANT: Feder, J. N.
; APPLICANT: Mintier, G.
; APPLICANT: Ramamathan, C. S.
; APPLICANT: Westphal, R.
; APPLICANT: Hawken, D. R.
; APPLICANT: Cacace, A.
; APPLICANT: Barber, L.
; APPLICANT: Kornacker, M. G.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY7,
; FILE REFERENCE: D0044NP
; CURRENT APPLICATION NUMBER: US/09/966,782A

```

```

; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,731
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/268,580
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/315,423
; PRIOR FILING DATE: 2001-08-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 370
; TYPE: PRT
; ORGANISM: RAT
US-09-966-782A-13

```

Query Match 16.5%; Score 355.5; DB 9; Length 370;

Best Local Similarity 29.9%; Pred. No. 1.4e-21; Matches 115; Conservative 65; Mismatches 133; Indels 71; Gaps 15;

```

QY 31 SPPEGLAVPTAVCLCFVVGSGN--VVTVM-----IGRYDMRTTNLYIGMAVSDL 85
DB 11 SPGSVGA-VAVPVI FALIFILGWVGNGLVAVLLQPGPSAMQERSTTDLFILNLAVADL 69
QY 86 -LILGLPFD--LYRLMRSRPWFGPLCR--LSLYVGECYATLHMTALSVERYLA 139
DB 70 CFLICVFPQAAIYTL--DAMLFGARVCKVHLIYL--TWYASSFTLAAVSDRYLA 123
QY 140 ICRPLRARVTVRRVRALIAVMAVALLSAGPFLVGVSDPGISVVPGLNGTARIAS 199
DB 124 VRHPLRRAALRTPRNANAVGLVLLAALFSAPYL-----SYGTVRGA 168
QY 200 SPLASSPPLWLSRAPSPSPGSETAEAAALFSRECRSPPAQLGARVMLWTTAYFPLP 259
DB 169 LELC--VPAM-----EDRRRALDVATF--AAGVILP 196
QY 260 FLCLSLYIGLIGELMSSRPLRGASGREGHQTKRVLLVVLATICMLPHVGR 319
DB 197 VAVVSLAYGRTLCFLMAAVGPAGAAAARARRATGRAGRAMLVVVVVGISMLPHHV-- 253
QY 320 IYINTEDSR--MMYFSQYFNIVALQLFYLSASINPILYMLISKYRAAFKLLARKSRP 377
DB 254 IHLMAEFGAPFLIPASFFERTLHCLAYSNSLNPLVYLSLASHFRRAFRLLPCGRRRH 313
QY 378 RGFHRSRDTAGEVAGDGTGVGY 401
DB 314 RHHRRAHRAALRVQPASSG--PAGY 336

```

RESULT 6

```

US-09-970-966-209
; Sequence 209, Application US/09970966
; Patent No. US20020173638A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Moles, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jianshun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-966-209

```

Query Match 16.2%; Score 349.5; DB 9; Length 453;

Best Local Similarity 28.5%; Pred. No. 5.7e-21; Matches 111; Conservative 64; Mismatches 170; Indels 45; Gaps 13;

QY 20 PALPCDERRC-----PPGALVPT--AVCLCLFVVGSGNVTV--MLGR 65
 Db 4 PSLPGSD---CSQIIDSHVPEFEVATWIKITLIVLLIFVWGLLGNATIRVTVLOK 60
 QY 66 YDMRTTMYLGSMASVLLI--LGLPDLR--LMSRRWVGGLCRSLVYEGCTY 123
 Db 61 KGYLOKEVTDHMYSLACSDILVFLIGMPFYSIIMPLTSSYTLSCKHTTIFECSTY 120
 QY 124 ATLLHMTALSVERYLAIICRPLARVLTTRRRVALLIYAVALLSAGPFLFVGEODP 183
 Db 121 ATLLHMTALSVERYLAIICRPLARVLTTRRRVALLIYAVALLSAGPFLFVGEODP 179
 QY 184 GISVYVGLNGTARLASSPLASPPMLSRAPPSPPSGPETAALFSECRSPAO-- 241
 Db 180 LVN-VPSHRLGLTCNRSSSTRHQQ-----PETSNNISCTNLSRMTVPOSS 223
 QY 242 -LGAALVMTATAYFPFLPCLSLYGLIGRELMSRRP--LRGPAASGRERGHQTKR 298
 Db 224 IFGAFVYVLLVLSVAFMCMNMVLMKSKOKSLAGTRPQLKSSSEBRTARQTI 283
 QY 299 VLLVVLAFIICWLPFHVGRITII--NTEDSRMYFSGYFNIVALQ--LFYLSASINPIL 354
 Db 284 FLRLIVVTLAVCMMPNQRIRIMAAKPKHDWTRSYFRAYMTLLPSETFYLLSVINPIL 343
 QY 355 YNLISKKYRAAFKLLIARKSRPRGFHRSR 384
 Db 344 YTVSSQCFRRVFOVLCCLSLQHANHEKR 373

RESULT 7
 US-09-825-294-209
 ; Sequence 209, Application US/09825294
 ; Patent No. US2002000491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Stol, John A.
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Fling, Steven P.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; FILE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.484C5
 ; CURRENT APPLICATION NUMBER: US/09/825.294
 ; NUMBER OF SEQ ID NOS: 215
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 209
 ; LENGTH: 453
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-825-294-209

Query Match 16.2%; Score 349.5; DB 10; Length 453;
 Best Local Similarity 28.5%; Pred. No. 5.7e-21;
 Matches 111; Conservative 64; Mismatches 170; Indels 45; Gaps 13;

QY 20 PALPCDERRC-----PPGALVPT--AVCLCLFVVGSGNVTV--MLGR 65
 Db 4 PSLPGSD---CSQIIDSHVPEFEVATWIKITLIVLLIFVWGLLGNATIRVTVLOK 60
 QY 66 YDMRTTMYLGSMASVLLI--LGLPDLR--LMSRRWVGGLCRSLVYEGCTY 123
 Db 61 KGYLOKEVTDHMYSLACSDILVFLIGMPFYSIIMPLTSSYTLSCKHTTIFECSTY 120
 QY 124 ATLLHMTALSVERYLAIICRPLARVLTTRRRVALLIYAVALLSAGPFLFVGEODP 183
 Db 121 ATLLHMTALSVERYLAIICRPLARVLTTRRRVALLIYAVALLSAGPFLFVGEODP 179
 QY 184 GISVYVGLNGTARLASSPLASPPMLSRAPPSPPSGPETAALFSECRSPAO-- 241
 Db 180 LVN-VPSHRLGLTCNRSSSTRHQQ-----PETSNNISCTNLSRMTVPOSS 223
 QY 242 -LGAALVMTATAYFPFLPCLSLYGLIGRELMSRRP--LRGPAASGRERGHQTKR 298

Db 224 IFGAFVYVLLVLSVAFMCMNMVLMKSKOKSLAGTRPQLKSSSEBRTARQTI 283
 QY 299 VLLVVLAFIICWLPFHVGRITII--NTEDSRMYFSGYFNIVALQ--LFYLSASINPIL 354
 Db 284 FLRLIVVTLAVCMMPNQRIRIMAAKPKHDWTRSYFRAYMTLLPSETFYLLSVINPIL 343
 QY 355 YNLISKKYRAAFKLLIARKSRPRGFHRSR 384
 Db 344 YTVSSQCFRRVFOVLCCLSLQHANHEKR 373

RESULT 8
 US-09-966-871-77
 ; Sequence 77, Application US/09966871
 ; Patent No. US20020127539A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kopin, Alan S.
 ; TITLE OF INVENTION: Assays for Identifying Receptors Having
 ; FILE OF INVENTION: Alterations in Signaling
 ; FILE REFERENCE: 00398/512002
 ; CURRENT APPLICATION NUMBER: US/09/966.871
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US 60/236,302
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/288,644
 ; NUMBER OF SEQ ID NOS: 87
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 77
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-966-871-77

Query Match 15.8%; Score 340; DB 10; Length 380;
 Best Local Similarity 25.5%; Pred. No. 2.7e-20;
 Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;

QY 6 NGSDBEGARREP---WPAALPCDERRCPPFLGALVPTAVCLCLFVVGSGNVTVML 62
 Db 39 NGVSSEDOQLEPAHISPAIP-----VITTAYSVVFVGLVGNLSLMPV 83
 QY 63 IGRYDMRTTMYLGSMASVLLI--LGLPDLR--LMSRRWVGGLCRSLVYEGCT 122
 Db 84 IIRYTKMTATNIYIFNLADALVTTMPFQ--SAVYLMNSMPFGVLCXIVISIDYNN 142
 QY 123 YATLLHMTALSVERYLAIICRPLARVLTTRRRVALLIYAVALLSAGPFLFVGEOD 182
 Db 143 FTSIFTLTMSVDRYIAVCHPVKALDFRPLKAKIINICIWILLASSVGSIAIVLGTKR 202
 QY 183 PGISVYVGLNGTARLASSPLASPPMLSRAPPSPPSGPETAALFSECRSPAO 242
 Db 203 EDVDVIE-----CSLOFPDDERSWMD-----LFWKIC----- 229
 QY 243 GALRYVLTATAYFPFLPCLSLYGLIGRELMSRRP--LRGPAASGRERGHQTKR 302
 Db 230 -VFVAFVIVPVLIIIVCYTLMRL--KSVRLSG--SREKDRMLRITIKIVL 278
 QY 303 VVLAFLIICWLPFHVGRITII--NTEDSRMYFSGYFNIVALQ--FYLSASINPIL 360
 Db 279 VVAVFIIICWLPFHVGRITII--NTEDSRMYFSGYFNIVALQ--FYLSASINPIL 335
 QY 361 KYRAA-----AFKLLIARKSRPRGFHRSR 396
 Db 336 NFKRCRDFCPPIKMERQST---NRVANYQDPAASRDVGG 375

RESULT 9
 US-10-039-645-77
 ; Sequence 77, Application US/10039645
 ; Patent No. US20020147170A1
 ; GENERAL INFORMATION:

APPLICANT: Kopin, Alan S.
 APPLICANT: Beinborn, Martin
 TITLE OF INVENTION: Constitutively Active, Hypersensitive,
 TITLE OF INVENTION: and No. US20020147170A1functional Receptors as No. US20020147170
 FILE REFERENCE: 00398/510002
 CURRENT APPLICATION NUMBER: US/10/039,645
 PRIOR FILING DATE: 2001-10-25
 PRIOR FILING DATE: 2000-10-26
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 77
 LENGTH: 380
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-039-645-77

Query Match 15.8%; Score 340; DB 12; Length 380;
 Best Local Similarity 25.5%; Pred. No. 2.7e-20;
 Matches 103; Conservative 75; Mismatches 146; Indels 80; Gaps 13;

QY 6 NGSDGEGAREPP--WPALPPCDERRCSPFLGALVPTAVCLCFVVGSGNVTVML 62
 DB 39 NGVSGEDQLEPAHISAP-----VITTAIVSVFVVGVLGNSLVMEV 83
 QY 63 IGRVDRRTTNLYGSAVSDLLILGLPDLRYLRSPVWFGELCRSLYVEGCT 122
 DB 84 IIRYTKMTATNIYIFNLADALVTTMPFO-SAVYLMNSWPFGLCKIVISIDYNNM 142
 QY 123 YATLHMTALSVRYLAICRPLARAVLTVRRVRALINLMAVALLSAPFLVVEOD 182
 DB 143 FTSFTLTMVVDRIACHVKAIDFRTPLAKATINICIMVLASVGISALVVGSTKR 202
 QY 183 PGISVVPGLNLTARIASPLASSPPLMLSRAPPSPGPTAEAAALFSPRCRPSAQL 242
 DB 203 EDVVDIE-----CSLQFPDDDEXWMD-----LFMKIC----- 229
 QY 243 GAAVMLVTTAYFPFLPCLISILYGLIGRELMSSRRPLRGPAASGRERGHQTKRVLLV 302
 DB 230 -----VFVFAFVLPVLIIVCYTLMILRL-KSVRLSG--SREKDRNLRITKIVLV 278
 QY 303 VLAFLICMLPEHGRITVI--NTEDSRMYFSQYENIVALOLFSLASINILYNLISK 360
 DB 279 VVAVFICMTPIHIFILVEALGSTSHSTAVSSYF--CIALGTNSLSLNEVLVAFIDE 335
 QY 361 KYRAA-----AFKLLARKSRPRGFHRSRDTAGEVAG--DTGG 396
 DB 336 NFKRCFRDFCPIMRMEROST---NRVNTVQDPASMRDVG 375

RESULT 10
 US-09-966-871-76
 Sequence 76, Application US/09966871
 Patent No. US20020127539A1
 GENERAL INFORMATION:
 APPLICANT: Kopin, Alan S.
 TITLE OF INVENTION: Assays for Identifying Receptors Having
 Alterations in Signaling
 FILE REFERENCE: 00398/512002
 CURRENT APPLICATION NUMBER: US/09/966,871
 CURRENT FILING DATE: 2001-09-28
 PRIOR FILING DATE: 2000-09-28
 PRIOR FILING DATE: 2000-09-28
 PRIOR APPLICATION NUMBER: US 60/243,550
 PRIOR FILING DATE: 2001-05-03
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 76
 LENGTH: 376
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-966-871-76

Query Match 15.6%; Score 336.5; DB 10; Length 376;
 Best Local Similarity 24.6%; Pred. No. 5.2e-20;
 Matches 102; Conservative 76; Mismatches 158; Indels 79; Gaps 12;

QY 1 MGSPWNGSDGEGAREPPWPAALP-----CDERRCSPFLGALVPV- 41
 DB 1 MESPFGEDEPTCA--PSACLPNNSAMFPGMAEPNSGASGEDAQLEPAHISPADPVE 57
 QY 42 -TAVCLCFVVGSGNVTVMLIGRYDMRTTNLYGSAVSDLLILGLPDLRYLMR 100
 DB 58 ITAVSVVFPVVGVLGNSLVMEVIRYTKMTATNIYIFNLADALVTTMPFO-STVYL 116
 QY 101 SRPWFGPLCRSLYVEGCTVATLHMTALSVRYLAICRPLARAVLTVRRVRALIA 160
 DB 117 MNSWPFEDVLCIVISIDYNNMFTSIFTLTMSVDRIACHVKAIDFRTPLAKATINI 176
 QY 161 VMAVALLSAGPFLVGVVEODPGISVVPGLNLTARIASPLASSPPLMLSRAPPSPPS 220
 DB 177 CIMLSSVGSISALVIGTVREDDVIE-----CSLQFPDDDEXWMD----- 219
 QY 221 GPPTAEAAALFSPRCRPSAQLGALRVMLVTTAYFPFLPCLISILYGLIGRELMSSRRP 280
 DB 220 -----LFMKIC-----VFVFAFVLPVLIIVCYTLMILRL-KSVRL 254
 QY 281 LRGPASGRERGHQTKRVLLVTVLAFIICMLPEHGRITVI--NTEDSRMYFSQYENI 338
 DB 255 LSG--SREKDRNLRITRVLVVAVFVCGMTPIHIFILVEALGSTSHSTAVSSYF-- 310
 QY 339 VALQFLYSINPILYNLISKYRAA-----AFKLLARKSRPRGFHRSRDTA 387
 DB 311 -CIALGTNSLSLNEVLVAFIDEENFKRCFRDFCPIMRMEROSTSRVNTVQDPA 364

RESULT 11
 US-10-039-645-76
 Sequence 76, Application US/10039645
 Patent No. US20020147170A1
 GENERAL INFORMATION:
 APPLICANT: Kopin, Alan S.
 APPLICANT: Beinborn, Martin
 TITLE OF INVENTION: Constitutively Active, Hypersensitive,
 TITLE OF INVENTION: and No. US20020147170A1functional Receptors as No. US2002014717
 FILE REFERENCE: 00398/510002
 CURRENT APPLICATION NUMBER: US/10/039,645
 CURRENT FILING DATE: 2001-10-25
 PRIOR FILING DATE: 2000-10-26
 PRIOR FILING DATE: 2000-10-26
 NUMBER OF SEQ ID NOS: 87
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 76
 LENGTH: 376
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-039-645-76

Query Match 15.6%; Score 336.5; DB 12; Length 376;
 Best Local Similarity 24.6%; Pred. No. 5.2e-20;
 Matches 102; Conservative 76; Mismatches 158; Indels 79; Gaps 12;

QY 1 MGSPWNGSDGEGAREPPWPAALP-----CDERRCSPFLGALVPV- 41
 DB 1 MESPFGEDEPTCA--PSACLPNNSAMFPGMAEPNSGASGEDAQLEPAHISPADPVE 57
 QY 42 -TAVCLCFVVGSGNVTVMLIGRYDMRTTNLYGSAVSDLLILGLPDLRYLMR 100
 DB 58 ITAVSVVFPVVGVLGNSLVMEVIRYTKMTATNIYIFNLADALVTTMPFO-STVYL 116
 QY 101 SRPWFGPLCRSLYVEGCTVATLHMTALSVRYLAICRPLARAVLTVRRVRALIA 160
 DB 117 MNSWPFEDVLCIVISIDYNNMFTSIFTLTMSVDRIACHVKAIDFRTPLAKATINI 176
 QY 161 VMAVALLSAGPFLVGVVEODPGISVVPGLNLTARIASPLASSPPLMLSRAPPSPPS 220

Db 177 CWMSSVGSIAVLGGTKREDVDVIE-----CSIQFPDDDDYSNMD----- 219
 QY 221 GPETAAALFSREGRSPPAQIALRWMLWTTAYFELPFLCLSIYGLIGRELMSSRRP 280
 Db 220 -----LFFKIC-----VFIFAFVPIPLIIVCYTLMILRL-KSVRL 254
 QY 281 LRGPASGRGRHQRKRVLLVVAFLICWLPFHGRIIYI--NEDSMWTFSCYFNI 338
 Db 255 LSG--SREKDRMLRITRLVLLVVAFFVVCWTPHIFILVEALGSISSHSTAALSIYF-- 310
 QY 339 VALQFYLASINPIILYNLSIKKYRAA-----AFKLLARKSRPFGHRSRDTA 387
 Db 311 -CIALGYNSSLNPIIYAFUDENFKCFRDFCPLKMRMRGOSTSRVRNTVQDPA 364

RESULT 12
 US-09-823-114-23
 ; Sequence 23, Application US/09823114
 ; Patent No. US20020061554A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EVANS, CHRISTOPHER J.
 ; KEITH, DUANE E.
 ; TITLE OF INVENTION: OPIOID RECEPTOR GENES
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & ROEKSTER
 ; STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1888
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/823,114
 ; FILING DATE: 29-Mar-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/148,351
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MORASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22000-20526.22
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030 MASNROERSMSH
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 367 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: Modified-site
 ; LOCATION: group(19, 26, 37)
 ; OTHER INFORMATION: /note="extracellular Asn residues
 ; that are consensus sites for N-linked glycosylation"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
 US-09-823-114-23

Query Match 15.6%; Score 335.5; DB 10; Length 367;
 Best Local Similarity 27.4%; Pred. No. 6.1e-20;
 Matches 101; Conservative 68; Mismatches 132; Indels 67; Gaps 11;

QY 34 PLGALVPTAVCLCFVVGSGNVVTVMILIGRYDMRTTNNYLGSMASVSDLLILGLPF 93
 Db 44 PLGKVTIVGLVAVCGILGNCLVMVYLHHTKTKATNIYIFPLADLTLLTLTF 103

QY 94 ---DLVRLMSRPVFGPLLCRLSLVYEGCTYATLLHMTALSVERYIATCPRLARVLV 150
 Db 104 QGTDLI-----LGPWFENALCKTVIAIDYNNMFTSTFLTMSVDRYVAICHPIRALDVR 159
 QY 151 TRRRVRALIAVMAVALLSAGPFLVGVEDPDGISVPGNGTARIASPLASPLML 210
 Db 160 TSSKAQAVNVAIVMALASV-----VGVPVAIMSAQVDEDEIEC-----L 198
 QY 211 SRAPPSPPSGPETAEAAALFSREGRSPPAQIALRWMLWTTAYFELPFLCLSIYGLI 270
 Db 199 VEIPTPDQWGPVPAICIFLS-----FIVPLVIVSCYSIM 235
 QY 271 GRELMSSRRPLRGPASGRGRHQRKRVLLVVAFLICWLPFHGRIIYINTDSMM 330
 Db 236 IRRLL-RGVRLLSG--SREKDRMLRITRLVLLVVAFFVVCWTPVQV---FVLAQGLGVQ 288
 QY 331 YFSQYFNIVALQF-----YLSASINPIILYNLSIKKYRAAFAKLLARKSR--PGFHRS 383
 Db 289 PSSS--TAVAILRFTCTALGVNSCLNPIIYAFUDENFKACFRKCCASALRDRVQSDRV 346
 QY 384 RDTAGEVA 391
 Db 347 RSLAKDVA 354

RESULT 13
 US-10-087-345A-15
 ; Sequence 15, Application US/10087345A
 ; Publication No. US20030045696A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ouyang, Chung
 ; TITLE OF INVENTION: ORPHANIN FQ RECEPTOR NUCLEIC ACIDS
 ; FILE REFERENCE: UM-06962
 ; CURRENT APPLICATION NUMBER: US/10/087,345A
 ; CURRENT FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 367
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-087-345A-15

Query Match 15.5%; Score 333.5; DB 9; Length 367;
 Best Local Similarity 26.8%; Pred. No. 8.9e-20;
 Matches 100; Conservative 68; Mismatches 132; Indels 73; Gaps 12;

QY 34 PLGALVPTAVCLCFVVGSGNVVTVMILIGRYDMRTTNNYLGSMASVSDLLILGLPF 93
 Db 44 PLGKVTIVGLVAVCGILGNCLVMVYLHHTKTKATNIYIFPLADLTLLTLTF 103
 QY 94 ---DLVRLMSRPVFGPLLCRLSLVYEGCTYATLLHMTALSVERYIATCPRLARVLV 150
 Db 104 QGTDLI-----LGPWFENALCKTVIAIDYNNMFTSTFLTMSVDRYVAICHPIRALDVR 159
 QY 151 TRRRVRALIAVMAVALLSAGPFLVGVEDPDGISVPGNGTARIASPLASPLML 210
 Db 160 TSSKAQAVNVAIVMALASV-----VGVPVAIMSAQVDEDEIEC-----L 198
 QY 211 SRAPPSPPSGPETAEAAALFSREGRSPPAQIALRWMLWTTAYFELPFLCLSIYGLI 270
 Db 199 VEIPTPDQWGPVPAICIFLS-----FIVPLVIVSCYSIM 235
 QY 271 GRELMSSRRPLRGPASGRGRHQRKRVLLVVAFLICWLPFHGRIIYINTDSMM 330
 Db 236 IRRLL-RGVRLLSG--SREKDRMLRITRLVLLVVAFFVVCWTPVQV---FVLAQGLGVQ 288
 QY 331 YFSQYFNIVALQF-----YLSASINPIILYNLSIKKYRAAFAKLLARKSRPGFHRS-- 383
 Db 289 PSSS--TAVAILRFTCTALGVNSCLNPIIYAFUDENFKACFRKCCASALRDRVQSDRV 346
 QY 384 RDTAGEVAGDTG 395

Db 342 VSDRVSTAKDVG 354

RESULT 14

US-09-992-331-16
Sequence 16, Application US/09992331
Publication No. US20030022186A1

GENERAL INFORMATION:

APPLICANT: FEDER, JOHN N.
APPLICANT: MINTIER, GABE
APPLICANT: RAMANATHAN, CHANDRA S.
APPLICANT: HAWKEN, DONALD R.
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBWMY18,
TITLE OF INVENTION: EXPRESSED HIGHLY IN PITUITARY GLAND AND COLON CARCINOMA
TITLE OF INVENTION: CELLS
FILE REFERENCE: D0048NP
CURRENT APPLICATION NUMBER: US/09/992,331
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/308,540
PRIOR FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: 60/261,782
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/248,483
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 428
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-992-331-16

Query Match 15.5%; Score 333.5; DB 9; Length 428;
Best Local Similarity 26.3%; Pred. No. 1.1e-19;
Matches 93; Conservative 64; Mismatches 129; Indels 67; Gaps 9;

QY 36 GAVPVTAVCCLFVVGSGNVVTVMILGRYKDMRTTNLYGSAVSDLLILGLPPL 95
Db 45 GILISLYLVVC--VGLGSLVLYVLRHTSSPSVTSVYLINALDELMLGLP-L 101
QY 96 YLRMSRPWVFGPLLRLSLVVGEGCTATLLHMTALSVERYLACRPLRARVLTTRRV 155
Db 102 AAGNALSYWPGSLMCRVMAVDGINOTSTFCILVMSVDRLAVVHPRKSARMTAPPA 161
QY 156 RALIAVMAVALLSAGPPLFLVGYEODPGISVVPGLNGTARIASSPLMLSRAP 215
Db 162 RMVSAAVWVASAVVVLVWVPSGV----- 185
QY 216 PSPPSGPEETAALFIRECR--PSPAQLGALRMVMTTAYFFLPFLCLSLYGLIGR 272
Db 186 ---PRGMS-----CHMOPPEPAAMRTAFITYTALGFPGPLVLCCLYLLIV 232
QY 273 ELWSSRRPLR-----GPAASGRERGRHROTKEVLVAVLAFICMLPFHVGRIITYNT 324
Db 233 KYRSTTRVAPSCQWQAPACQRRRSERVRVTVMAVVALFLVCMPIFYLANIVNVVC 292
QY 325 E-DSKMYEQYENIVALQLFYLSASINPILYNTLSKKYRAAFAKLLIARSR 376
Db 293 PLDEEPAFGLVFLVVALP--YANSCANPILYGLISYRFK--QGFRILLRPSR 342

RESULT 15

US-09-966-782A-16
Sequence 16, Application US/09966782A
Publication No. US20030022183A1

GENERAL INFORMATION:

APPLICANT: Battaglin, P.
APPLICANT: Feder, J. N.
APPLICANT: Mintier, G.
APPLICANT: Ramanathan, C. S.
APPLICANT: Westphal, R.
APPLICANT: Hawken, D. R.
APPLICANT: Cacace, A.

APPLICANT: Barber, L.
APPLICANT: Kornacker, M. G.
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBWMY7,
TITLE OF INVENTION: EXPRESSED HIGHLY IN SPINAL CORD
FILE REFERENCE: D0044NP
CURRENT APPLICATION NUMBER: US/09/966,782A
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/235,731
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/268,580
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/315,423
PRIOR FILING DATE: 2001-08-28
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 372
TYPE: PRT
ORGANISM: RAT
US-09-966-782A-16

Query Match 15.4%; Score 331; DB 9; Length 372;
Best Local Similarity 27.4%; Pred. No. 1.4e-19;
Matches 110; Conservative 62; Mismatches 133; Indels 96; Gaps 16;

QY 6 NSDGPGEARP---PWPALPCDERRCSPPLGALVPTAVCLCFVVGSGNVVTVM 61
Db 4 SSGQGAENTSGEGSGGQW-----PEAVLVPL--FPALIFVGTGNAVLVA 48
QY 62 LIGRYDMRTTNLYGSAVSDLLILGLPPL--LYLRMSRPWVFGPLLRLSLDYVG 118
Db 49 VLIRGQAVSTNLLFTINLGVADLCEILCCVPOATITVL--DDWVFGSLCKAVHFLI 105
QY 119 ECGTYATLLHMTALSVERYLACRPLRARVLTTRRVRALIIVMAVALLSAGPFLVIG 178
Db 106 FLTMHSSFTLAIVSDRLYLAIRYPLHSRELRTPRNALAIIGLWGLALFSGGYLSYR 165
QY 179 VEODPGISVVPGLNGTARIASSPLMLSRAPPPSPSGPEETAALFIRECRPS 238
Db 166 OSQLANLV-----CHPAW-----SAPRR----- 184
QY 239 PAQGLARVMTTAY--FPLPFLCLSLYGLIGREIWSRRPLRGPAASGRERGRHRTK 297
Db 185 -----RAMDCTFVESYLLPVLVLTATRTLRVLRVTVDPV--TAGSGSGQAKRKVT 235
QY 298 RVLVAVLAFICMLPEHV-----GRIIYINTEDSRMYEQYENIVALQLFYLSASI 350
Db 236 RMIIIVAVLFCICMMDHALLILCWFGRPPL-----TRATY--ALRISHLVYANSCV 287
QY 351 NPILYNTLSKKYRAAFAKLLIARSRPGRFHRSDTAGYVA 391
Db 288 NPIVVALVSKHFR-KGFRKICAGLKP-----APRRASGRVS 323

Search completed: March 16, 2003, 15:08:53
Job time: 16 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:13 ; Search time 17 seconds

(without alignments)
713,073 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156

Sequence: 1 MGSPWNGSDGPEGAREPPWP.....DTGGDTGYTETSANVKTWG 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	857.5	39.8	366	4	US-09-077-675A-13
2	857	39.7	353	4	US-09-077-675A-3
3	855.5	39.7	364	4	US-09-077-675A-16
4	854	39.6	361	4	US-09-077-675A-8
5	769.5	35.7	302	4	US-09-077-675A-2
6	768.5	35.6	302	4	US-09-077-675A-7
7	668.5	31.0	271	4	US-09-077-675A-12
8	638.5	29.6	289	4	US-09-077-675A-10
9	627	29.1	289	4	US-09-077-675A-5
10	477	22.1	415	4	US-09-545-944-2
11	463	21.5	353	1	US-08-118-270-45
12	463	21.5	353	5	PCT-US93-08528-45
13	407	18.9	410	3	US-08-858-876A-2
14	407	18.9	410	3	US-09-472-880-2
15	401	18.6	416	3	US-08-858-876A-4
16	401	18.6	416	4	US-09-472-880-4
17	391.5	18.2	410	4	US-09-200-090-2
18	382	17.7	391	4	US-09-200-090-4
19	358.5	16.6	319	3	US-08-832-339-2
20	358.5	16.6	319	4	US-09-372-438-2
21	356.5	16.5	418	1	US-07-816-283-10
22	356.5	16.5	418	1	US-08-417-103-10
23	338	15.7	380	4	US-08-168-275A-5
24	338	15.7	380	4	US-09-351-198-5
25	338	15.7	380	4	US-09-113-426-5
26	335.5	15.6	367	3	US-08-676-351-2
27	335.5	15.6	367	4	US-08-405-271A-23

28	335	15.5	352	4	US-09-029-027B-2	Sequence 2, Appl
29	334	15.5	380	4	US-08-765-743-2	Sequence 2, Appl
30	333.5	15.5	367	2	US-08-454-548-2	Sequence 2, Appl
31	333.5	15.5	367	3	US-08-454-552-2	Sequence 2, Appl
32	333.5	15.5	367	3	US-08-147-592A-6	Sequence 6, Appl
33	333.5	15.5	367	3	US-08-889-108-17	Sequence 17, Appl
34	333.5	15.5	367	4	US-08-293-694A-6	Sequence 6, Appl
35	333.5	15.5	367	5	PCT-US94-10358-17	Sequence 17, Appl
36	333	15.4	380	3	US-08-676-351-5	Sequence 5, Appl
37	331	15.4	372	2	US-08-626-685A-8	Sequence 8, Appl
38	331	15.4	372	4	US-08-993-088A-2	Sequence 2, Appl
39	331	15.4	372	4	US-08-993-424B-2	Sequence 2, Appl
40	331	15.4	372	4	US-08-665-034A-4	Sequence 4, Appl
41	330.5	15.3	367	3	US-09-170-331-4	Sequence 4, Appl
42	330.5	15.3	428	1	US-07-816-283-12	Sequence 12, Appl
43	330.5	15.3	428	1	US-08-417-103-12	Sequence 12, Appl
44	330	15.3	376	4	US-08-387-707-17	Sequence 17, Appl
45	330	15.3	376	4	US-08-405-271A-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-077-675A-13
Sequence 13, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ

COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A

FILING DATE: 3-JUN-1998
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720

TELEX:
INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
LENGTH: 366 amino acids

TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-077-675A-13

Query Match 39.8%; Score 857.5; DB 4; Length 366;
Best Local Similarity 44.5%; Pred. No. 1,5e-65;

Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

QY 5 WNGSDPEGA-----REPPWALPPCD---ERRCSEPPGALVPTAVCLCFVGVSGN 56
 Db 2 WNAFSEBPGFNLTLADLDMDASPGNDSLGDELLQLFPAPLAGVATCATVAFVVGIAAG 61

QY 57 VVTVMILGRYRDMRTTNNIYLSMAVSDLLILGLPEFDLYLRMRSPWVFGPLCRSLY 116
 Db 62 LITMLVSRFRRLRTTNNIYLSMAFSDLLIFCMLPLDLRLMQRPWNGDLCKLFGP 121

QY 117 VGEQCTVATLHMTALSVRYLAICRPLARVLTTRRRVALLIIVMAVALSAGEPLFL 176
 Db 122 VSECTVATVLTITLSTVRYPALICPLRAKVVTKGRVLTIVFVIMAVAFSAGPIFVL 181

QY 177 VGEQDPGISVVGINGTARIASSPPLMLSRAPPSPPSGPETAALFSPRECR 236
 Db 182 VGVEHE-----NGT-----DP-W-----DTNECR 199

QY 237 PS-PAQGLALVMLVMTAVFPLPLCLILXGLIGRELMSRRPLRGPASGERGHR 294
 Db 200 PTEFAVRSGLLTVMWVSSIFFLPVFCLTVLSLIGRKLRRRGDAVVGASLRDQNHK 259

QY 295 QCRVLLVVLVLAIFICMLPFHVGRITVINTEDS---RMVFSQYFNIVALQFLYSASIN 351
 Db 260 QIVKMLAVVFAFILLICMLPFHVGRYLFPSKSFEPGSLIAQISQYCNLVSVFLYLSAIN 319

QY 352 PLYNLISSKRYRAAFKLLIARKSRPRGFRSDTAGEVAGDTGDTGVGTETSAN 407
 Db 320 PLYNIMSKRYRAVAFRLGLFEPFSQKSLTKDESSR-----AWTESSIN 365

RESULT 2
 US-09-077-675A-3
 ; Sequence 3, Application US/09077675A
 ; Patent No. 6242199
 ; GENERAL INFORMATION:
 ; APPLICANT: Pai, Lee-Yuh
 ; APPLICANT: Feighner, Scott C.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Pong, Sheng-Shung
 ; APPLICANT: Van Der Ploeg, Leonardus H.T.
 ; TITLE OF INVENTION: RECEPTOR ASSAY
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/077, 675A
 ; FILING DATE: 3-JUN-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cocuzzo, Anna L.
 ; REGISTRATION NUMBER: 42,452
 ; REFERENCE/DOCKET NUMBER: 19590P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 732-594-1273
 ; TELEFAX: 732-594-4720
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 353 amino acids

TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-077-675A-3

Query Match 39.7%; Score 857; DB 4; Length 353;
 Best Local Similarity 46.2%; Pred. No. 1,6e-65;
 Matches 185; Conservative 53; Mismatches 96; Indels 66; Gaps 9;

QY 19 WPAIPPCD---ERRCSEPPGALVPTAVCLCFVGVSGNVVVMILGRYRDMRTTNNL 75
 Db 8 WDAPEPNDLSVEBELPLFPPLLAGVATCATVAFVVGIAAGLITMLVSRFRRLRTTNNL 67

QY 76 YLSMAVSDLLILGLPEFDLYLRMRSPWVFGPLCRSLYVGEQCTVATLHMTALSVE 135
 Db 68 YLSMAFSDLLIFCMLPLDLRLMQRPWNGDLCKLFGVSECTVATVLTITALSVE 127

QY 136 RYLAICRPLARVLTTRRRVALLIIVMAVALSAGEPLFLVGVGEQDPGISVVGINGTA 195
 Db 128 RYFALICPLRAKVVTKGRVLTIVIMAVAFSAGPIFVLVGVEHD-----NGT- 177

QY 196 RIASSPPLMLSRAPPSPPSGPETAALFSPRECRPS-PAQGLALVMLVMTAVTT 253
 Db 178 -----DPRD-----TWECATERAVASGLITVMWVSS 205

QY 254 AYFPLPCLILXGLIGRELMSRRPLRGPASG---REGRHROTQKRVLLVVLAFIIC 310
 Db 206 VFFLPVFCLTVLSLIGRKLW---RRKRGDAVGSRLRDQNHKQIVKMLAVVFAFILLIC 262

QY 311 WLPFHVGRITV---INTEDSRMTFSQYFNIVALQFLYSASINPLIYNLISSKRYRAAF 367
 Db 263 WLPFHVGRYLFPSKSFEPGSLIAQISQYCNLVSVFLYLSAINPLIYNLISSKRYRAAF 322

QY 368 KULLIARKSRPRGFRSDTAGEVAGDTGDTGVGTETSAN 407
 Db 323 KLLGPEPFSQKSLTKDESSR-----AWTESSIN 352

RESULT 3
 US-09-077-675A-16
 ; Sequence 16, Application US/09077675A
 ; Patent No. 6242199
 ; GENERAL INFORMATION:
 ; APPLICANT: Pai, Lee-Yuh
 ; APPLICANT: Feighner, Scott C.
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Pong, Sheng-Shung
 ; APPLICANT: Van Der Ploeg, Leonardus H.T.
 ; TITLE OF INVENTION: RECEPTOR ASSAY
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Merck & Co., Inc.
 ; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
 ; CITY: Rahway
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07065-0900
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/077, 675A
 ; FILING DATE: 3-JUN-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cocuzzo, Anna L.
 ; REGISTRATION NUMBER: 42,452

```

; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-16

Query Match          39.7%; Score 855.5; DB 4; Length 364;
Best Local Similarity 48.0%; Pred. No. 2.2e-65;
Matches 184; Conservative 52; Mismatches 82; Indels 65; Gaps 11;

QY 5 WNGSDGPEGARER-----FWPALPPCD---ERRCSFPLGALVPTAVCLCFVVGVS 55
DB 2 WNAAT--PSEEPENVTLDLDWDASPGNDSLPDELLPFPAPLAGVTATCAVLFVVGISG 59

QY 56 NVVTWALIRYRDMRTTNNLYLGSMVAVSDLLILGLPFDLYRLMRSPWVFGPLICRLSL 115
DB 60 NLTMLVVRFRFELRTTNLYLSSMAFSDLLIFLCMPDLVRLMQYRPMWFGDLCFLFQ 119

QY 116 YVGECTVATLLHMTLASVERYLACRPLRVLVTRRRVRLALVAVALLSAGPFLF 175
DB 120 FVSECTVATVLTTLTASVERYPALCEPLRAKVVTGKRVKVLIVAVAFSAGPIFV 179

QY 176 LVGVQDPGISVVGNGTARLASPLASPLMLSRAPPPSGGELTAAALFSREC 235
DB 180 LVGVHE-----NGT-----DPRD-----TNEC 197

QY 236 RPS--PAQGLALRWLMTVTAFFLPCLSLYLGLIGRELMSRRPLRGPAASG---RE 290
DB 198 RATEFAVNSGLLTVMWVSVFFELVFCLTVYSLGRKLM--RR--KDDAVAGSLRD 253

QY 291 RGHROTQRLVLLVLAFLICMLPFHVGRITTYINTEDS---RMWYFSQYENIVALQLFYLS 347
DB 254 QNHKQTVKMLAVVPAFILLCMPLFHVGRYLFKSPFGSLIAQISQYCNLVSPVLFYLS 313

QY 348 ASINPIILNLSKKYRAAFKLL 370
DB 314 AAINPIILNLSKKYRAVAFKLL 336

RESULT 4
US-09-077-675A-8
; Sequence 8, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTA for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077, 675A

```

```

; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 361 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-8

Query Match          39.6%; Score 854; DB 4; Length 361;
Best Local Similarity 45.8%; Pred. No. 2.9e-65;
Matches 182; Conservative 57; Mismatches 98; Indels 60; Gaps 8;

QY 19 WPALPPCD---ERRCSFPLGALVPTAVCLCFVVGVSANVTMLGRYRDMRTTNL 75
DB 16 WDASPGNDSLGDELLIQFPPALGAVTATCAVLFVVGISGNTLMLVVRFRFELRTTNL 75

QY 76 YLGSMVAVSDLLILGLPFDLYRLMRSPWVFGPLICRLSLYVGECTVATLLHMTLASVE 135
DB 76 YLSSMAFSDLLIFLCMPDLVRLMQYRPMWFGDLCFLFQFVSECTVATVLTALASVE 135

QY 136 RYLAICRPLRVLVTRRRVRLALVAVALLSAGPFLVGVQDPGISVVGNGTAR 195
DB 136 RYPAICRPLRVLVTRRRVRLALVAVALLSAGPFLVGVHE-----NGT- 185

QY 196 RIASSPPLMLSRAPPPSGGELTAAALFSRECPS--PAQGLALRWLMTV 253
DB 186 -----DB--W-----DTNECPTERAVNSGLLTVMWVSS 213

QY 254 AYFPLPCLSLYLGLIGRELMSRRPLRGPAASGERGHROTQRLVLLVLAFLICMP 313
DB 214 IFPLPVPFCLTVYSLGRKLMRRRGDAVVGASLRDQNHQTVKMLAVVPAFILLCMPL 273

QY 314 FHVGRITTYINTEDS---RMWYFSQYENIVALQLFYLSASINPIILNLSKKYRAAFKLL 370
DB 274 FHVGRYLFKSPFGSLIAQISQYCNLVSPVLFYLSAINPIILNLSKKYRAVAFKLL 333

QY 371 LARKSRPGRFHRSDTAGEVAGDTGDPVGYTFSAN 407
DB 334 GFEPFSQRKSLTKDESSR-----AWTESSIN 360

RESULT 5
US-09-077-675A-2
; Sequence 2, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA

```

ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-2

Query Match 35.7%; Score 769.5; DB 4; Length 302;
Best Local Similarity 46.3%; Pred. No. 3,8e-58;
Matches 163; Conservative 48; Mismatches 80; Indels 63; Gaps 8;

QY 60 VMLIGRYDMRTTNTLYLGSMAVSDLLILGLPDLRYLRMRSPWFGLRLSLTYGE 119
DB 1 MLVVSFRFRLTNTLYLSSMAFSDLLIFLCMPDLRLMQRPNWNGLLCKLFGFVSE 60
QY 120 GCTVATLTLHMTALSVRYLACRPLRVRVLTBRVRALIAVMAVALSAGPFLVGV 179
DB 61 SCTVATVLTTLASVRYLACRPLRVRVLTBRVRALIAVMAVALSAGPFLVGV 120
QY 180 BODPGISVPGINGTARLASSPLASPPMLSRAPPPSPGPETAEAALFSRECRPS- 238
DB 121 EHD-----NGT-----DPRD-----TNECRATE 138
QY 239 -PAQGLALVWLWTTAVYFPLPCLSLIYLIGRELMSSRRPLRGPAAG--REGRHR 294
DB 139 FAVRSGLLTVMWVWSVFFLPVCLTVLSLIGRKLW---RRRKGAAVSSLRDQNHK 195
QY 295 QTRVLLVWVLAFLICMLPEHVGRIIY---INTEDSRMWYFSQYFNIALQFYLSASIN 351
DB 196 QTVMAVAVVFAFLICMLPEHVGRIYLSKSLPQSVETIAQISQYCNLVSVFLFYLSAIN 255
QY 352 PLYNLISKYRAAFAKLLIARKSRPRGFRHSRDTAGEVADTGDTGVYTETSAN 407
DB 256 PLYNLISKYRAVAVFRLIGPEPFSQRLSTLKDESSR-----AWTESSIN 301

RESULT 6
US-09-077-675A-7
Sequence 7, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.

STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-7

Query Match 35.6%; Score 768.5; DB 4; Length 302;
Best Local Similarity 46.2%; Pred. No. 4.7e-58;
Matches 163; Conservative 51; Mismatches 82; Indels 57; Gaps 7;

QY 60 VMLIGRYDMRTTNTLYLGSMAVSDLLILGLPDLRYLRMRSPWFGLRLSLTYGE 119
DB 1 MLVVSFRFRLTNTLYLSSMAFSDLLIFLCMPDLRLMQRPNWNGLLCKLFGFVSE 60
QY 120 GCTVATLTLHMTALSVRYLACRPLRVRVLTBRVRALIAVMAVALSAGPFLVGV 179
DB 61 SCTVATVLTTLASVRYLACRPLRVRVLTBRVRALIAVMAVALSAGPFLVGV 120
QY 180 BODPGISVPGINGTARLASSPLASPPMLSRAPPPSPGPETAEAALFSRECRPS- 238
DB 121 EHD-----NGT-----DP--W-----DTNECRPTE 138
QY 239 -PAQGLALVWLWTTAVYFPLPCLSLIYLIGRELMSSRRPLRGPAAGREGRHRQTK 297
DB 139 FAVRSGLLTVMWVWSVFFLPVCLTVLSLIGRKLWRRRDAVAGSLRDQNHQTV 198
QY 298 RVLVWVLAFLICMLPEHVGRIIYINTEDS---RMWYFSQYFNIALQFYLSASINPIL 354
DB 199 KMLAVVAVFALICMLPEHVGRIYLSKSLPQSVETIAQISQYCNLVSVFLFYLSAINPIL 258
QY 355 YNLISKYRAAFAKLLIARKSRPRGFRHSRDTAGEVADTGDTGVYTETSAN 407
DB 259 YNLISKYRAVAVFRLIGPEPFSQRLSTLKDESSR-----AWTESSIN 301

RESULT 7
US-09-077-675A-12
Sequence 12, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.

```

1  TITLE OF INVENTION: RECEPTOR ASSAY
2  NUMBER OF SEQUENCES: 16
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE: Merck & Co., Inc.
5  STREET: P.O. Box 2000, 126 E. Lincoln Ave.
6  CITY: Rahway
7  STATE: NJ
8  COUNTRY: USA
9  ZIP: 07065-0900
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Diskette
12 COMPUTER: IBM Compatible
13 OPERATING SYSTEM: DOS
14 SOFTWARE: PASTESEQ for Windows Version 2.0
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/09/077,675A
17 FILING DATE: 3-JUN-1998
18 CLASSIFICATION:
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER:
21 FILING DATE:
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Cocuzzo, Anna L.
24 REGISTRATION NUMBER: 42,452
25 REFERENCE/DOCKET NUMBER: 19590P
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: 732-594-1273
28 TELEFAX: 732-594-4720
29 TELEX:
30 INFORMATION FOR SEQ ID NO: 12:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 271 amino acids
33 TYPE: amino acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: protein
37 US-09-077-675A-12

Query Match 31.0%; Score 668.5; DB 4; Length 271;
Best Local Similarity 44.7%; Pred. No. 1,4e-49;
Matches 144; Conservative 44; Mismatches 77; Indels 57; Gaps 7;

QY 91 LPFDLYRLMRSPWVGPLLCRLSLVVGEGCTATLHMTALVERLACRPLRAVLV 150
DB 1 MFLDLVRLMQYPMNFGDLCKLFOVSBSCTATVLTALVERFAICPLRAKVV 60
QY 151 TRRVRLIAVMAVALLSAGPFLFLVGVGDPGISVPGINGTARIASSPLASPL 210
DB 61 TKGRVGLVFIWMAVAFCSAGPIFVLVGEHE-----NGT-----DP--W- 98
QY 211 SRAPPSPSPSGETAAALFSRECRPS--PAOLGALRWMLVTTTAVFPLPCLSL 268
DB 99 -----DINECRPTIEFAVSSGLITVWVSSIFFFLPVFCLTVLYS 138
QY 269 LIGRELWSSRRPLRGPAASGRGRHQRKVLVVVLAFLICWLPVHVGRIIINEDS- 327
DB 139 LIGRLMRRRREDAVGASLRDQNKQTKMLAVVFAFLICLPHVGVYLSKSPFEPG 198
QY 328 --RMVFSQYENIVALQLEYLSASINPILYNTISKYRAAFAKLLARKSRPGFHSRD 385
DB 199 SLEIAQISQYCNLVSVFLFYLSAIPILYNTIMSKYRVAVFRLLPFEPSPQKSLTKD 258
QY 386 TAGEVAGDGTGDTGVGYTETSAN 407
DB 259 ESSR-----AMTSSIN 270

RESULT 8
US-09-077-675A-10
; Sequence 10, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pat, Lee-Yuh
```

```

1  APPLICANT: Feigner, Scott C.
2  APPLICANT: Howard, Andrew D.
3  APPLICANT: Pong, Sheng-Shung
4  APPLICANT: Van Der Ploeg, Leonardus H.T.
5  TITLE OF INVENTION: RECEPTOR ASSAY
6  NUMBER OF SEQUENCES: 16
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: Merck & Co., Inc.
9  STREET: P.O. Box 2000, 126 E. Lincoln Ave.
10 CITY: Rahway
11 STATE: NJ
12 COUNTRY: USA
13 ZIP: 07065-0900
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: PASTESEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/077,675A
21 FILING DATE: 3-JUN-1998
22 CLASSIFICATION:
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER:
25 FILING DATE:
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Cocuzzo, Anna L.
28 REGISTRATION NUMBER: 42,452
29 REFERENCE/DOCKET NUMBER: 19590P
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 732-594-1273
32 TELEFAX: 732-594-4720
33 TELEX:
34 INFORMATION FOR SEQ ID NO: 10:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 289 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 US-09-077-675A-10

Query Match 29.6%; Score 638.5; DB 4; Length 289;
Best Local Similarity 43.6%; Pred. No. 5.4e-47;
Matches 143; Conservative 45; Mismatches 79; Indels 61; Gaps 9;

QY 5 WNGSDSGEGA-----REPPPALPPCD--BRGSPFLGALVPTAVCLCFVGVSGN 56
DB 2 WNATPSEBPGNLTLDLDDMDASPGNDSLGDELLOLFPAPLAGVTATCVALFVVGIA 61
QY 57 VVTVMILGRYADMTTNTNLVLSMAVSDLLILGLFPDLYRLMRSPWVFGPLCRLSLY 116
DB 62 LITMLVSRPRELRTTNTNLVLSMAVSDLLILGLFPDLYRLMRSPWVFGPLCRLSLY 121
QY 117 VGBSCTATLHMTALSVERYLACRPLRAVLVTRRRVRLIAVMAVALLSAGPFLFL 176
DB 122 VBSCTATVLTALSVERYFAICPLRAKVVTKGVKLVFIWMAVAFCSAGPIFVL 181
QY 177 VGVGODGISVVPGLNTATARIASSPLASPLWLSRAPSPSPSGETAAALFSRECR 236
DB 182 VGVGEHE-----NGT-----DP--W-----DINECR 199
QY 237 PS--PAOLGALRWMLVTTTAVFPLPCLSLYGLIGRELWSSRRPLRGPAASGRGRH 294
DB 200 PTEFAVSSGLITVWVSSIFFFLPVFCLTVLYSLIGRLMRRRREGDAVAGASLRDQ 259
QY 295 QTKRVL-----LVVLA---FITCMLP 313
DB 260 QTKVMKGSGQALRLSLAGPIISLCILP 287

RESULT 9
US-09-077-675A-5
```

```

; Sequence 5, Application US/0907675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feigener, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P. O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-077-675A-5

Query Match      29.1%; Score 627; DB 4; Length 289;
Best Local Similarity 44.3%; Pred. No. 5, 1e-46;
Matches 117; Conservative 43; Mismatches 71; Indels 58; Gaps 8;

```

5 WNSDDEGA-----REPPALPPCD---ERRCSPPFLGALVPTAVCLCFVGVSGN 56
2 WNAIPSEEPGNLTLPLGMDAPPENDSLVELPLFFPTLLAGTATCAVLFVGVIGN 61
57 VVTVMILGRYDKMTTNNLYLSMAVSDLLILGLPFDLYIRMSRPVVFGLCRSLY 116
62 LITLVVSRFRKMTTNNLYLSMAFSELLIFLCWPLFLFRLMOYRPNMLGLCKLRF 121
117 VVGSCYATLTLMTALSTERYLATCRPLRAVLYTRRRVRLIYVLMVALISAGPLFL 176
122 VSECTATVTLTALSTERYFAICFPLRAKVVTGKVKLVIVMAVAFCSAGPIFVL 181
177 VGVQDQGISVVPGLNGTARIASSPLASSPPLMLSRAPPPSPGPETAAMALFSRRCR 236
182 VGVHND-----NGT-----DPRD-----TNECR 199
237 PS--PAOLGALVWLWTTAFLPFLCLSLIYGLIGRELMSRRPLKPPASG---RER 291
200 ATEPAVNSGLLTVMVWSVFEFLFVFCITVLYSLIGRLT---RRKRGAAVGSRLDQ 256
292 GHRQTKVL 300
257 NHRQTVKML 265

```

RESULT 10
US-09-545-944-2
; Sequence 2, Application US/09545944
; Patent No. 6461836
; GENERAL INFORMATION:
; APPLICANT: AMES, ROBERT
; APPLICANT: ELSHOURBAGY, NABIL
; APPLICANT: MICHALOVICH, DAVID
; APPLICANT: SARAU, HENRY
; APPLICANT: SHABON, USMAN
; APPLICANT: VAMIER, LISA
; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
; TITLE OF INVENTION: (AXOR34) AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP70657-1
; CURRENT APPLICATION NUMBER: US/09/545,944
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 09/435,384
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-545-944-2

Query Match      22.1%; Score 477; DB 4; Length 415;
Best Local Similarity 30.5%; Pred. No. 4, 8e-33;
Matches 113; Conservative 76; Mismatches 102; Indels 78; Gaps 10;

```

30 CSPPLGALVPTAVCLCFVGVSGNVVYVWMLGRYDMRTTNLYLSMAVSDLLI-L 88
37 CGPRSRHFFLPSVYVYVIFVGVIGNVLCVLIHQAMKPTNYVFLSLAVSDLVLL 96
89 IGLPFDLYIRMSRPVVFGLCRSLYVSGCTATLMTALSTERYLATCRPLRARV 148
97 LGMPLVEYEMRNYPFLGVPVCYKTKALFETVVCASISITTVSEYRVAILHPRAL 156
149 LVTRRRVRLIYVLMVALISAGPLFLVGVV--ODPGISVVPGLNGTARIASSPLASSP 206
157 QSTRRAIRILGIWGSFVLPSPNTSIHGIRFHPNGLVPG-----SATCTVIK 208
207 PLWLSRAPPPSPGPETAAMALFSRRCRSPDQGLARVWLWTTAVFF-LPFLCLSI 265
209 PMWI-----YNFTIQVTSFLFLPMTVISY 234
266 LYGLIGRELMSR-----RPLRGPAASGRGRGRTQKRVLLVLAFLICWLP 313
235 LYLMLALRKDKSLAEADGNANIQRPCK-----KAVNKMFLVLLVLAFLICWLP 283
314 FVHGRITVINTED--SRMWFQYFNIVYALQFYLSASINPIYNIISKYRAAAEKLL 371
264 FHIDLFSPVEWETESLAAVFNLVHVGSLFYISSAVNPITNLSRFQ--AARQNTI 342
372 ARKSRPRGFR 382
343 S-----SFHK 347

```

RESULT 11
US-08-118-270-45
; Sequence 45, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK

```

```

; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-45

Query Match      21.5%; Score 463; DB 1; Length 353;
Best Local Similarity 32.3%; Pred. No. 6.2e-32;
Matches 122; Conservative 55; Mismatches 111; Indels 90; Gaps 11;

QY 39 VPTAVVCLCFVGVGSGVNVVYTMIGR--YRDMRTTNLYLGSMAVSDLLILGLPDL 95
DB 2 VLVTAIYALFLFVGVGNSVTAFTLARKKSLQSLQSTVHYHLSLSLSDLLILMW--EL 59
QY 96 YRLRSRPWFVFGPLLCRLSLYVGEQCTYATLLHMTLSVERYLAIICRPLRARIIVTRRV 155
DB 60 YNFMHMHMAFGDAGCRGYFLPADCTATATLVASLSVERYLAIICHPFAKTLMSRSRT 119
QY 156 RALIAVLAVALLSAGPFLFVGVGQDPGISVVPGLNGTARIASSPLASSPPLMLSRAP 215
DB 120 KKFISAIWLASALLAIPLMFTLGLQNRSGDGTGHPG--GLVCTPIVDT----- 164
QY 216 PSPSPGPTAAALFSPRECRSPAQOLGALRVMVTTAYFFL--PFLCLSLYGLIGREL 274
DB 165 -----ATVKVVIQVNTFMSFLPMLVISILNTVANKL 197
QY 275 WSSRRPLRGPAASGR-----ERGHROTQR-----VLVVVLAFLIIC 311
DB 198 TWV---VHQAAEGRCVCTVGTNGHLEHSTFMNRIGPRVQALRHGVLVRAVVAIFVVCW 254
QY 312 LPPHVGRIIYINTEDSRMYFS--QYFNIVALQLFYLASINPILYNLISKRYRAAFKL 369
DB 255 LPY---LCYISDEQWRTFLDFYHYFYMLTNALFYVSSAINPILYNLVSANFRQVFLST 310
QY 370 -----LLARKSRP 377
DB 311 LACLFCGWPILIRKKRP 328

RESULT 12
PCT-US93-08528-45
; Sequence 45, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF

```

```

; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08528-45

Query Match      21.5%; Score 463; DB 5; Length 353;
Best Local Similarity 32.3%; Pred. No. 6.2e-32;
Matches 122; Conservative 55; Mismatches 111; Indels 90; Gaps 11;

QY 39 VPTAVVCLCFVGVGSGVNVVYTMIGR--YRDMRTTNLYLGSMAVSDLLILGLPDL 95
DB 2 VLVTAIYALFLFVGVGNSVTAFTLARKKSLQSLQSTVHYHLSLSLSDLLILMW--EL 59
QY 96 YRLRSRPWFVFGPLLCRLSLYVGEQCTYATLLHMTLSVERYLAIICRPLRARIIVTRRV 155
DB 60 YNFMHMHMAFGDAGCRGYFLPADCTATATLVASLSVERYLAIICHPFAKTLMSRSRT 119
QY 156 RALIAVLAVALLSAGPFLFVGVGQDPGISVVPGLNGTARIASSPLASSPPLMLSRAP 215
DB 120 KKFISAIWLASALLAIPLMFTLGLQNRSGDGTGHPG--GLVCTPIVDT----- 164
QY 216 PSPSPGPTAAALFSPRECRSPAQOLGALRVMVTTAYFFL--PFLCLSLYGLIGREL 274
DB 165 -----ATVKVVIQVNTFMSFLPMLVISILNTVANKL 197
QY 275 WSSRRPLRGPAASGR-----ERGHROTQR-----VLVVVLAFLIIC 311
DB 198 TWV---VHQAAEGRCVCTVGTNGHLEHSTFMNRIGPRVQALRHGVLVRAVVAIFVVCW 254
QY 312 LPPHVGRIIYINTEDSRMYFS--QYFNIVALQLFYLASINPILYNLISKRYRAAFKL 369
DB 255 LPY---LCYISDEQWRTFLDFYHYFYMLTNALFYVSSAINPILYNLVSANFRQVFLST 310
QY 370 -----LLARKSRP 377
DB 311 LACLFCGWPILIRKKRP 328

RESULT 13
US-08-858-876A-2
; Sequence 2, Application US/0858876A
; Patent No. 6022856

```



```

; GENERAL INFORMATION:
; APPLICANT: Daniel CAPUT
; APPLICANT: Pascale CHALON
; APPLICANT: Pascual FERRARA
; APPLICANT: Vita NATALIO
; TITLE OF INVENTION: Type 2 Neurotensin Receptor
; TITLE OF INVENTION: (hnt-R2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,876A
; FILING DATE: 19-SEP-1997
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 9723204
; FILING DATE: 17-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,049
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-858-876A-2

```

```

Query Match      18.9% Score 407; DB 3; Length 410;
Best Local Similarity 31.2% Pred. No. 4.4e-27;
Matches 129; Conservative 60; Mismatches 135; Indels 90; Gaps 16;

QY 17 PPPPALP---CDERRCPPLGALVPTAVCLCFVVGSGNVTVMLI-----GR 65
DB 8 PPPSSPGLSLDARLGVDTRLMAKVLFTALYALIMAGAGNALSVHVLKARAGR 67
QY 66 YRDMRTTNLYGSMANVDLLIL-GLPFDLYR-LMRSRPWVGPIRLRLSLYVGEQCTY 123
DB 68 LRH-----HVLSTLALAGLLLLVGVPELVLSFVWFHYPWFGDLGCRGYFVHELCA 120
QY 124 ATLHMTALSVRYLAICPLRARVLTTRRRVRLAIIVLMAVALLSAGPFLVGV---- 179
DB 121 ATVLVAGLSARCLAVCPPLARSLTPRRTRMLVALSMAASLGLALPMVAVIMQCKHEL 180
QY 180 -----EODPGISVPGINGTARIASS-----PLASSPPL-----WLS 211
DB 181 ETADGEPBPASRCVTL--VSRTALQVFIQVNVLVSVLPALTAFNGVTVSHLLALCS 238
QY 212 RAPPPSPGPGETAEMAAALFSRCRSPPAOLGALRWMLVTTTAVFPLPCLSLIYGLIG 271
DB 239 QVPSTSTP-GSSTPSRLLELSEE-----GLLSFIWVKKT-----FIQG--G 276
QY 272 RELMSRRPLRGPAASGERGHROTGRVLLVVLAFIICWLPFHVGRITTYINTEDSR-- 328
DB 277 QVSLVHRHVDV-----RIRSLQRSVQVLRALIVMIVICWLPFHARLMTCYVDDAMTD 330
QY 329 -MMYFSQYFNIVALQFYLSASINPIYVNLISKYRAAAKXLLARKSRRGPH 381
DB 331 PLVNFYHYFVMTNTLFYVSSAVTPLLXNAVSSFR-----KLFLKAVSSSLCGEH 380

```

RESULT 14
US-09-472-880-2

```

; Sequence 2, Application US/09472880
; Patent No. 6274333
; GENERAL INFORMATION:
; APPLICANT: Daniel CAPUT
; APPLICANT: Pascale CHALON
; APPLICANT: Pascual FERRARA
; APPLICANT: Vita NATALIO
; TITLE OF INVENTION: Type 2 Neurotensin Receptor
; TITLE OF INVENTION: (hnt-R2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/472,880
; FILING DATE: 28-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 9723204
; FILING DATE: 17-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,049
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-09-472-880-2

```

```

Query Match      18.9% Score 407; DB 4; Length 410;
Best Local Similarity 31.2% Pred. No. 4.4e-27;
Matches 129; Conservative 60; Mismatches 135; Indels 90; Gaps 16;

QY 17 PPPPALP---CDERRCPPLGALVPTAVCLCFVVGSGNVTVMLI-----GR 65
DB 8 PPPSSPGLSLDARLGVDTRLMAKVLFTALYALIMAGAGNALSVHVLKARAGR 67
QY 66 YRDMRTTNLYGSMANVDLLIL-GLPFDLYR-LMRSRPWVGPIRLRLSLYVGEQCTY 123
DB 68 LRH-----HVLSTLALAGLLLLVGVPELVLSFVWFHYPWFGDLGCRGYFVHELCA 120
QY 124 ATLHMTALSVRYLAICPLRARVLTTRRRVRLAIIVLMAVALLSAGPFLVGV---- 179
DB 121 ATVLVAGLSARCLAVCPPLARSLTPRRTRMLVALSMAASLGLALPMVAVIMQCKHEL 180
QY 180 -----EODPGISVPGINGTARIASS-----PLASSPPL-----WLS 211
DB 181 ETADGEPBPASRCVTL--VSRTALQVFIQVNVLVSVLPALTAFNGVTVSHLLALCS 238
QY 212 RAPPPSPGPGETAEMAAALFSRCRSPPAOLGALRWMLVTTTAVFPLPCLSLIYGLIG 271
DB 239 QVPSTSTP-GSSTPSRLLELSEE-----GLLSFIWVKKT-----FIQG--G 276
QY 272 RELMSRRPLRGPAASGERGHROTGRVLLVVLAFIICWLPFHVGRITTYINTEDSR-- 328
DB 277 QVSLVHRHVDV-----RIRSLQRSVQVLRALIVMIVICWLPFHARLMTCYVDDAMTD 330
QY 329 -MMYFSQYFNIVALQFYLSASINPIYVNLISKYRAAAKXLLARKSRRGPH 381
DB 331 PLVNFYHYFVMTNTLFYVSSAVTPLLXNAVSSFR-----KLFLKAVSSSLCGEH 380

```

RESULT 15
US-08-858-876A-4Sequence 4, Application US/08858876A
Patent No. 6022856GENERAL INFORMATION:
APPLICANT: Daniel Caput

APPLICANT: Pascale CHALON

APPLICANT: Pascale FERRARA

APPLICANT: Vito NATALIO

TITLE OF INVENTION: Type 2 Neurotensin Receptor

TITLE OF INVENTION: (hmt-R2)

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC

STREET: 400 Seventh Street

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/858,876A

FILING DATE: 19-SEP-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR 9723204

FILING DATE: 17-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Player, William E.

REGISTRATION NUMBER: 31,049

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 416 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-858-876A-4

Query Match 18.6%; Score 401; DB 3; Length 416;
Best Local Similarity 29.4%; Pred. No. 1.5e-26;

Matches 122; Conservative 68; Mismatches 145; Indels 80; Gaps 13;

QY 18 PWPALP-----PCDERRCSPPLGALVPTAVCLCFVVGSGNVVTVMLIGRYDMRT 71

DB 6 PWPRLPSPAGLSLEBARLGVDTRLMAKVLFALYSLIFAFGTAGNALSHVVIKARAGR 65

QY 72 -TTNLVYLSMAVSDLLILL-GLPDLR-LMRSRPVWFGPPLCRSLVYEGCTYATLH 128

DB 66 GLRYHVHVLISLALLLLVSMFELYNFVWSHYVWFGDLGCRGYFVRELCAVATVLS 125

QY 129 MTALSVERYLAICRPLRVRVLTTRRRVALLVMAVALLSAGPFLVGV----- 179

DB 126 VASLSAEKCLAVCCPLRRRLITPRTRRLSLVWASLGALPMAVIMGQKHEVESADG 185

QY 180 EQDPGISVPELNGTARIA-----SSPLASSPPLMLSRAPPSPSGPETAABA 228

DB 186 EPEPASRVCTVLRATLQVFIVNVVSPALPLALTAFL-----NGITVNHLM 234

QY 229 ALFERECRSPAPOLGALVMTAVTAYFFPLCLSLVGLIGRELMSRRPLRGPASG 288

DB 235 ALYS-QVPSAQVSSIFSRLELSE-----EGLIGFITWRKTLISLGVASLV 281

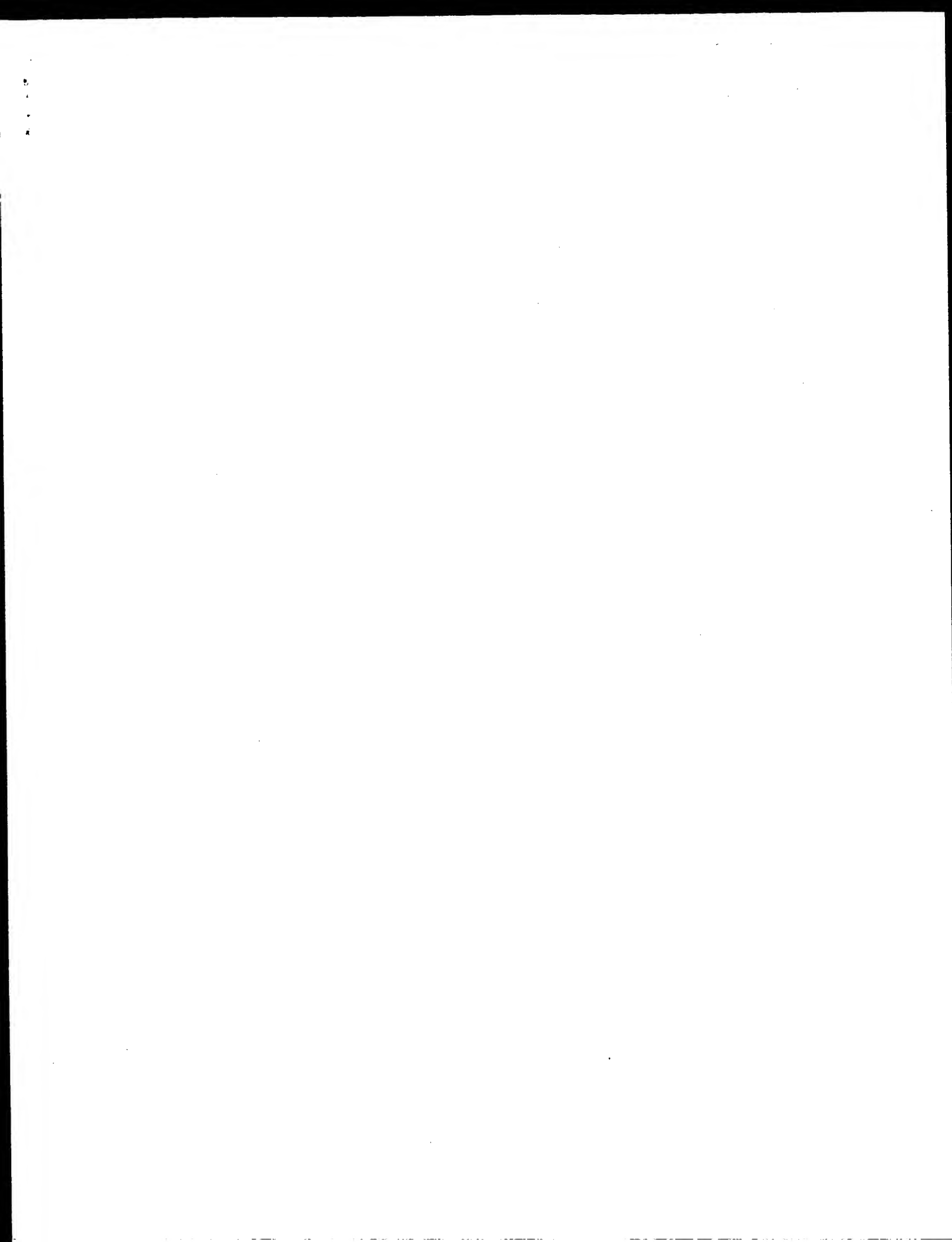
QY 289 RERGHROT-----RVLLVVVLAIIICMLPFGVRIIYINTD-----SRMYFSQYENI 338

DB 282 RHKDAQIRSLQHSQAQVRAIVAVVVICMLPYHARLMVCIYIPDQWINEIDFYHYFTW 341

QY 339 VALQFLYLSASINILVNLISKYRAAFAKLLARKSRPRGFHRSRDTAGEVAGD 393

DB 342 VTNLFFYVSSAVTPIIXNAVSSFR-----KLFL-----ESTGLGCE 379

Search completed: March 16, 2003, 15:05:54
Job time : 20 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 16, 2003, 15:03:07 ; Search time 40 Seconds

(without alignment)
1372.481 Million cell updates/sec

Title: US-09-876-252-130

Perfect score: 2156
Sequence: 1 MGSPWNGSDPEGAREPWP.....DTGDTVTGTTTSANKTNG 412

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SIDSI/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2156	100.0	412	21	AA02854
2	2149	99.7	412	21	AA02854
3	2149	99.7	412	22	AA02852
4	2149	99.7	412	22	AA02847
5	2149	99.7	501	22	AA05822
6	1741	80.8	400	22	AA08477
7	1575	73.1	386	21	AA04146
8	1575	73.1	386	22	AA02653
9	1143.5	53.0	271	22	AA08476
10	901	41.8	363	21	AA054147

11	901	41.8	363	22	AA08479	Amino acid sequenc
12	864.5	40.1	366	21	AA09066	Human mutant G pro
13	859.5	39.9	349	21	AA06293	A canine growth ho
14	857.5	39.8	366	21	AA09063	Human G protein-co
15	857.5	39.8	366	21	AA07034	Human G protein-co
16	857.5	39.8	366	22	AA07376	Rat growth hormone
17	857.5	39.8	366	22	AA06260	Human G-protein co
18	857	39.7	353	18	AA01921	Pig growth hormone
19	857	39.7	353	18	AA01921	Swine growth hormo
20	855.5	39.7	364	21	AA05455	A mouse growth hor
21	855.5	39.6	364	22	AA09377	Rat growth hormone
22	854	39.6	362	18	AA01921	Human growth hormo
23	854	39.6	362	18	AA01921	Human growth hormo
24	851.5	39.5	364	18	AA01921	Rat growth hormone
25	839.5	38.9	364	18	AA01921	Rat growth hormone
26	668.5	31.0	271	18	AA01921	Human growth hormo
27	668.5	31.0	271	18	AA01921	Human growth hormo
28	638.5	29.6	289	18	AA01921	Human growth hormo
29	635	29.5	289	18	AA01921	Human growth hormo
30	635	29.5	289	18	AA01921	Human growth hormo
31	633.5	29.4	289	18	AA01921	Human growth hormo
32	498	23.1	418	22	AA06371	Non-endogenous hum
33	495	23.0	418	17	AA08562	Human neurotensin
34	491	22.8	542	22	AA01231	Novel human diagno
35	490.5	22.8	403	21	AA09067	Human mutant G pro
36	485.5	22.5	445	22	AA06852	Human novel cytoxi
37	485.5	22.5	445	22	AA06856	Human novel cytoxi
38	484.5	22.5	403	21	AA09063	Human G protein-co
39	484.5	22.5	403	21	AA04464	Human growth hormo
40	484.5	22.5	403	22	AA08915	Human FM-3. Homo
41	484.5	22.5	426	22	AA03628	Human G-protein co
42	481	22.3	415	21	AA05292	Human neurotensin-
43	481	22.3	415	22	AA06336	Amino acid sequenc
44	477	22.1	412	22	AA06780	Amino acid sequenc
45	477	22.1	415	22	AA06780	Amino acid sequenc

ALIGNMENTS

RESULT 1	AA02854	AA02854 standard; Protein; 412 AA.
ID	AA02854	
XX	AA02854	
AC	AA02854	
XX	AA02854	
DT	22-AUG-2000	(first entry)
DE	Human G protein coupled receptor hGPR38 (V297K) protein SEQ ID NO:130.	
KW	Human; G protein coupled receptor; GPCR; transmembrane receptor;	
KW	identification; agonist; screening; therapeutic; pharmaceutical;	
KW	mutant.	
XX		
OS	Homo sapiens.	
XX	Synthetic.	
XX		
PN	WO200022131-A2.	
PD	20-APR-2000.	
XX		
PF	13-OCT-1999;	99WO-US24065.
XX		
XX	13-OCT-1999;	98US-0170496.
PR	12-NOV-1998;	98US-0108029.
PR	20-NOV-1998;	98US-0109213.
PR	27-NOV-1998;	98US-0110060.
PR	16-FEB-1999;	99US-0120416.
PR	26-FEB-1999;	99US-0121852.
PR	12-MAR-1999;	99US-0123944.
PR	12-MAR-1999;	99US-0123945.
PR	12-MAR-1999;	99US-0123946.
PR	12-MAR-1999;	99US-0123948.

PR 12-MAR-1999; 99US-0123949.
 PR 12-MAR-1999; 99US-0123951.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 28-MAY-1999; 99US-0137567.
 PR 30-JUN-1999; 99US-0141448.
 PR 27-AUG-1999; 99US-0151114.
 PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156655.
 PR 29-SEP-1999; 99US-0156634.
 (AREN-) ARENA PHARM INC.

Behan DP, Lehmann-Brunema K, Chalmers DT, Chen R, Dang HT,
 Gore M, Law CM, Lin I, Lowitz K, White C,
 WPI: 2000-317986/27.
 N-PSDB; AAA46116.

Non-endogenous, human G protein-coupled receptors for screening
 receptor, inverse or partial agonists useful as therapeutic agents -
 Example 2; Page 168-169; 187pp; English.

The present invention describes transmembrane receptors, preferably
 human G protein coupled receptors (GPCR), for which the endogenous
 ligand is unknown (orphan GPCR receptors). More specifically the present
 invention relates to non-endogenous, constitutively activated versions
 of a human GPCR. These non-endogenous human GPCRs can be useful for
 the direct identification of candidate compounds as receptors agonists,
 inverse agonists or partial agonists for use as pharmaceutical agents,
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 the exemplification of the present invention.

Sequence 412 AA;

Query Match 100.0%; Score 2156; DB 21; Length 412;
 Best Local Similarity 100.0%; Pred. No. 4.8e-195;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSFPMNGSDGEGARBPMPALPCDERCSPPLGALVPTAVCLCFVVGSGNVTV 60
 DB 1 MGSFPMNGSDGEGARBPMPALPCDERCSPPLGALVPTAVCLCFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYGSAVSDLLILGLPFDLYRLMRSPWVFGPLCRSLYVGBG 120
 DB 61 MLIGRYDMRTTNNLYGSAVSDLLILGLPFDLYRLMRSPWVFGPLCRSLYVGBG 120
 QY 121 CTYATLHMTALSVRYIAICRPLARVLYRRRRAIIVIMANALLSAGFLVNGE 180
 DB 121 CTYATLHMTALSVRYIAICRPLARVLYRRRRAIIVIMANALLSAGFLVNGE 180
 QY 181 ODPISVVGANGTARIASPPLASPPMLSRAPPPSPSGETAAALFSGRCPSPA 240
 DB 181 ODPISVVGANGTARIASPPLASPPMLSRAPPPSPSGETAAALFSGRCPSPA 240
 QY 241 QLGALRWLVWTTAVFPLPCLSLIYGLIGRELWSSRRPLGPAASGERGHRQTKRVL 300
 DB 241 QLGALRWLVWTTAVFPLPCLSLIYGLIGRELWSSRRPLGPAASGERGHRQTKRVL 300
 QY 301 LVVTLAFICMLPRVGGIITYINTEDSRMYRQYENTVAOLFLYISASINPLNLISK 360
 DB 301 LVVTLAFICMLPRVGGIITYINTEDSRMYRQYENTVAOLFLYISASINPLNLISK 360
 QY 361 KYRAAFAFLIARSRPRGFRHSRDTAGVADTGDTGDTGYETISANVKTWG 412
 DB 361 KYRAAFAFLIARSRPRGFRHSRDTAGVADTGDTGDTGYETISANVKTWG 412

RESULT 2
 AAY54145
 ID AAY54145 standard; Protein; 412 AA.
 AC AAY54145;
 XX
 DT 27-MAR-2000 (first entry)
 XX

Amino acid sequence of the motilin receptor splice variant MTL-R1A.

XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation.

XX Homo sapiens.
 OS
 XX
 PN W09964436-A1.
 XX
 PD 16-DEC-1999.
 XX
 PF 08-JUN-1999; 99WO-US12773.
 XX
 PR 12-JUN-1998; 98US-0089098.
 XX

(MERI) MERCK & CO INC.

Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
 Pong S, Smith RG;
 WPI: 2000-105868/09.
 N-PSDB; AA245403.

Novel receptor protein for screening compounds used in treating
 irritable bowel syndrome, constipation and other gastric conditions -
 Claim 3; Fig 3; 44pp; English.

XX The present sequence represents splice variant MTL-R1A of the motilin
 CC receptor. The gene encodes a G-protein coupled receptor, and is
 CC designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,
 CC MTL-R1A (see AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a
 CC functional seven transmembrane domain form, and MTL-R1B is a truncated
 CC five transmembrane domain. The MTL-R1 proteins are used to identify
 CC agonists and antagonists which can be used for treating gastric motility
 CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.

Sequence 412 AA;

Query Match 99.7%; Score 2149; DB 21; Length 412;
 Best Local Similarity 99.8%; Pred. No. 2.2e-194;
 Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSFPMNGSDGEGARBPMPALPCDERCSPPLGALVPTAVCLCFVVGSGNVTV 60
 DB 1 MGSFPMNGSDGEGARBPMPALPCDERCSPPLGALVPTAVCLCFVVGSGNVTV 60
 QY 61 MLIGRYDMRTTNNLYGSAVSDLLILGLPFDLYRLMRSPWVFGPLCRSLYVGBG 120
 DB 61 MLIGRYDMRTTNNLYGSAVSDLLILGLPFDLYRLMRSPWVFGPLCRSLYVGBG 120

Db 61 MLIGRYDMRTTNLYIGSMAVSDLLILGLPFDLYRLMSRPWFPGPLLCLSLIYVSEG 120
 QY 121 CTYATLLHMTALSVERYLAICRPLRARVLYTRRRVRALIAVLMAVALLSAGPFLVGE 180
 Db 121 CTYATLLHMTALSVERYLAICRPLRARVLYTRRRVRALIAVLMAVALLSAGPFLVGE 180
 QY 181 QDPGISVVPGLNGTARLASSPLASSPPLMSRAPPPSPGPGPTAEAAALFSRECRPSPA 240
 Db 181 QDPGISVVPGLNGTARLASSPLASSPPLMSRAPPPSPGPGPTAEAAALFSRECRPSPA 240
 QY 241 QLGALRWMLMTTAYFPLPFLCLSLIYGLIGRELMSRRPLRGPAASGRERGHROTAVL 300
 Db 241 QLGALRWMLMTTAYFPLPFLCLSLIYGLIGRELMSRRPLRGPAASGRERGHROTAVL 300
 QY 301 LVVLAFLIICMLPFIHVGRITTYINTEDSRMWFSSQYFNIVALOLFYLASINPILYNLSK 360
 Db 301 LVVLAFLIICMLPFIHVGRITTYINTEDSRMWFSSQYFNIVALOLFYLASINPILYNLSK 360
 QY 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGEVAGDTGDTVGTYETISANKYTMG 412
 Db 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGEVAGDTGDTVGTYETISANKYTMG 412

RESULT 3

AAB62652
 ID AAB62652 standard; Protein; 412 AA.

AC AAB62652;

DT 23-JUL-2001 (first entry)

DE Long form of motilin receptor, GPR-38A isoform.

KM zsig33; signal transduction; hormone; enzyme; neural development;
 KM gastric contractility; nutrient uptake; digestive; pancreatic; human;
 KM insulin-like growth factor-1; growth hormone; bone; gastrointestinal;
 KM glucose; osteopathic; anorectic; vulnery; immunomodulator; GHS-R;
 KM G-protein coupled receptor; motilin receptor; GPR38; GPR-38A; isoform.

OS Homo sapiens.

XX WO200138355-A2.

XX 31-MAY-2001.

PF 22-NOV-2000; 2000MO-US32074.

PR 22-NOV-1999; 99US-0166765.

XX (ZYMO) ZYMOGENETICS INC.

PI Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;

DR WPI; 2001-355879/37.

DR N-PSDB; AAF83683.

PT Forming reversible peptide receptor complex for purifying cell and
 PT peptides, stimulating signal transduction and modulating hormone
 PT secretion, involves contacting a receptor with zsig33 polypeptide
 XX
 PS Disclosure; Page 102-104; 111pp; English.

CC The invention relates to a method of forming a reversible peptide-
 CC receptor complex that involves providing an immobilized receptor, and
 CC contacting the receptor with a zsig33 peptide (comprising residues 24-37
 CC of AAB62649), where the receptor binds to the zsig33 peptide. The method
 CC is useful for purifying cells, purifying a peptide, stimulating signal
 CC transduction in a cell expressing a receptor. It is also useful for
 CC modulating secretion of hormones, neural development and/or utilization,
 CC gastric contractility, nutrient uptake, secretion of digestive and
 CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
 CC -1, secretion of non-zsig33 proteins. It is useful for modulating growth
 CC hormone secretion in a mammal having a disease associated with abnormal

CC levels of growth hormone, such as osteoporosis, bone repair, bone
 CC remodeling, low osteoblast levels, cartilage repair and remodeling,
 CC skeletal dysplasia, immune suppression, obesity, growth retardation,
 CC protein catabolic responses after surgery, cachexia, protein loss,
 CC dwarfism, wound healing and ovulation induction, treating a mammal having
 CC a metabolic disorder requiring neurological feedback, such as satiety
 CC regulation, glucose absorption and metabolism and neuropathy-associated
 CC gastrointestinal disorders, and stimulating glucose-induced insulin
 CC release in a mammal. The present sequence represents the long form of
 CC motilin receptor, GPR-38A (one of the two isoforms of GPR38 which result
 CC from alternative splicing). GPR38 has homology to the human G-protein
 CC coupled receptor, GHS-R.

XX SQ Sequence 412 AA;

Query Match 99.7%; Score 2149; DB 22; Length 412;

Best Local Similarity 99.8%; Pred. No. 2.2e-194;

Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGAREPPWPALPCDERRCSPPLGALVPTAVCLCFVVGSGNVTV 60

Db 1 MGSPWNGSDGEGAREPPWPALPCDERRCSPPLGALVPTAVCLCFVVGSGNVTV 60

QY 61 MLIGRYDMRTTNLYIGSMAVSDLLILGLPFDLYRLMSRPWFPGPLLCLSLIYVSEG 120

Db 61 MLIGRYDMRTTNLYIGSMAVSDLLILGLPFDLYRLMSRPWFPGPLLCLSLIYVSEG 120

QY 121 CTYATLLHMTALSVERYLAICRPLRARVLYTRRRVRALIAVLMAVALLSAGPFLVGE 180

Db 121 CTYATLLHMTALSVERYLAICRPLRARVLYTRRRVRALIAVLMAVALLSAGPFLVGE 180

QY 181 QDPGISVVPGLNGTARLASSPLASSPPLMSRAPPPSPGPGPTAEAAALFSRECRPSPA 240

Db 181 QDPGISVVPGLNGTARLASSPLASSPPLMSRAPPPSPGPGPTAEAAALFSRECRPSPA 240

QY 241 QLGALRWMLMTTAYFPLPFLCLSLIYGLIGRELMSRRPLRGPAASGRERGHROTAVL 300

Db 241 QLGALRWMLMTTAYFPLPFLCLSLIYGLIGRELMSRRPLRGPAASGRERGHROTAVL 300

QY 301 LVVLAFLIICMLPFIHVGRITTYINTEDSRMWFSSQYFNIVALOLFYLASINPILYNLSK 360

Db 301 LVVLAFLIICMLPFIHVGRITTYINTEDSRMWFSSQYFNIVALOLFYLASINPILYNLSK 360

QY 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGEVAGDTGDTVGTYETISANKYTMG 412

Db 361 KYRAAFAFKLLARKSRPRGFHRSRDTAGEVAGDTGDTVGTYETISANKYTMG 412

RESULT 4

AAB68478

XX AAB68478 standard; Protein; 412 AA.

AC AAB68478;

DT 23-JUL-2001 (first entry)

DE Amino acid sequence of a human motilin receptor polypeptide.

XX Motilin receptor; gastrointestinal disease; gastric motility disorder;

XX Gastro paresis; irritable bowel syndrome; diarrhoea.

XX Homo sapiens.

XX WO200132710-A1.

XX 10-MAY-2001.

XX 25-OCT-2000; 2000MO-US29426.

XX 29-OCT-1999; 99US-0162264.

XX (MERI) MERCK & CO INC.

ID	AAB68477
XX	AAB68477 standard; Protein; 400 AA.
AC	AAB68477;
DT	23-JUL-2001 (first entry)
DE	Amino acid sequence of a rabbit motilin receptor polypeptide.
KM	Motilin receptor; gastrointestinal disease; gastric motility disorder; gastroparesis; irritable bowel syndrome; diarrhoea.
OS	Oryctolagus cuniculus.
PN	MO200132710-A1.
PD	10-MAY-2001.
PF	25-OCT-2000; 200OWO-US29426.
PR	29-OCT-1999; 99US-0162264.
PA	(MERI) MERCK & CO INC.
PI	Tan C, McKee K;
DR	WPI; 2001-343479/36. N-PEDB; AAF85448.
XX	
PT	Novel polypeptides related to dog and rabbit motilin receptor polypeptide, comprising unique regions from dog and motilin receptor amino acid sequence, useful for identifying compounds for treating diarrhoea in humans _
XX	
PS	Claim 8; Page 18; 42pp; English.
CC	The present sequence represents a rabbit motilin receptor polypeptide. The specification describes an unique sequence present in exon 1 of the dog motilin receptor, which is not present in human or Sphaeroides nephelus 75E7 motilin receptor sequences. The unique nucleic acid sequence is useful for measuring the ability of a compound to affect motilin receptor activity. Motilin receptor polynucleotides and peptides are used to identify therapeutic compounds which are useful for treating gastrointestinal diseases and disorders such as gastric motility disorders, gastroparesis, irritable bowel syndrome, and diarrhoea.
SQ	Sequence 400 AA;
Query Match	80.8%; Score 1741; DB 22; Length 400;
Best Local Similarity	83.7%; Pred No. 7, 1e-156;
Matches 343; Conservative 11; Mismatches 44; Indels 12; Gaps 3	
Y	1 MGSPNWGSDGPGEAREPPWPALPCDERCSPPLGALVTVAVCLCFVGVSGNVTV 60
D	1 MGSPWNCGSDGPEDAREPMAALPCDERCSPFLGTIVPVATVAGLGFVGVSNNVT 60
Y	61 MLIGGYRMRTTNTNYLGSMAYSDLLILGLFPDYLRLMRSPWFGPLLKSLYVGG 120
D	61 LLIGGYRDMRTTNTNYLGSMAYSDLLILGLFPDYLRLMRSPWFGQLCRSLYVGG 120
Y	121 CTYATLLHMTALVERLYLAICRPLARVLVTRRRVRALTAVMAYVALSAGFPLVGV 180
D	121 CTYASLLHMTALSVRYLAICRPLARVLVTRRRVRALIAIMAYVALSAGFFFLVGV 180
Y	181 ODPGISVTPGLNGTARIASSPLASSPEMLWSRAPPPSPSGETAAMALLFSRECRPS 240
D	181 QDPAAFAAPDRNGTV-----PLDPSSP-----AASPSPGC-GAAMAALLFSRECRPS 228
Y	241 QLGAIRVWLWTATAEPFLPCLSTLYGLIGELWSSRPPLRGSPAASGERGHROTQRTL 300
D	229 QLGLIRVWLWTATAEPFLPCLSTLYGLIARQLWRGRPLRGPAATGERGHROTQRTL 288

Qy	301	LVVLAFAIIICMIPFHVGHIIITYINTEDSSMMVFESQYFNIVAAQLFPLASINPIIYNLISK	360
Db	289	LVVLAFAIICMIPFHVGHIIITYINTEDSSMMVFESQYFNIVAAQLFPLASINPIIYNLISK	348
Qy	361	KYRAAAFKILLARKSRPRGFHRSRDTAGEVADTGDTGVGTETGSANYKT	410
Db	349	KYRAAARFLLRESRAGPSGVCSSRGPEQVADTGDTAGCTETSANTKT	398
RESULT 7			
ID	AAV54146		
XX	AAV54146 standard; Protein; 386 AA.		
XX	AAV54146;		
XX	27-MAR-2000 (first entry)		
DE	Amino acid sequence of the motilin receptor splice variant MTL-R1B.		
KW	Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;		
KW	spliced form; MTL-R1A; MTL-R1B; gastric motility disorder;		
KW	functional defect; neurological disorder; scleroderma; colonoscopy;		
KW	paraneoplastic syndrome; radiation induced dysmotility; diabetes;		
KW	infection; stress-related motility disorder; psychogenic disorder;		
KW	gastroparasis; gastro-oesophageal reflux disease; constipation;		
KW	chronic idiopathic pseudo obstruction; acute faecal impaction;		
KW	postoperative ileus; gallstones; infantile colic; diarrhoea;		
KW	irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain;		
XX	endoscopy; duodenal intubation.		
OS	Homo sapiens.		
PN	WO9964436-A1.		
XX	16-DEC-1999.		
PD	08-JUN-1999; 99WC-US12773.		
PF	12-JUN-1998; 98US-0089098.		
FR	(MERI) MERCK & CO INC.		
PA	Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;		
XX	Pong S, Smith RG;		
PI	WPI; 2000-105868/09.		
DR	N-PSDB; AAZ45404.		
XX	Novel receptor protein for screening compounds used in treating		
PT	irritable bowel syndrome, constipation and other gastric conditions		
PS	Claim 5; Fig 5; 44pp; English.		
XX	The present sequence represents splice variant MTL-R1B of the motilin		
CC	receptor. The gene encodes a G-protein coupled receptor, and is		
CC	designated MTL-R1 (also GPR38). Two spliced forms of GPR38 exist,		
CC	MTL-R1A (see AAV54145) and MTL-R1B (see AAV54146). MTL-R1A is a		
CC	functional seven transmembrane domain form, and MTL-R1B is a truncated		
CC	five transmembrane domain. The MTL-R1 proteins are used to identify		
CC	agonists and antagonists which can be used for treating gastric motility		
CC	disorders, functional defects, disorders secondary to neurological		
CC	disorders e.g. scleroderma, paraneoplastic syndromes radiation induced		
CC	dysmotility, diabetes, infections, stress-related motility disorders,		
CC	psychogenic disorders, gastroparasis, gastro-oesophageal reflux disease,		
CC	constipation, chronic idiopathic pseudo obstruction, acute faecal		
CC	impaction, postoperative ileus, gallstones, infantile colic, irritable		
CC	bowel syndrome, non-ulcer dyspepsion, non-cardiac chest pain and		
CC	diarrhoea. They can also be used in the preparation for colonoscopy,		
CC	endoscopy and duodenal intubation.		
XX	Sequence 386 AA;		

Query Match 73.1%; Score 1575; DB 21; Length 386;

Best Local Similarity 99.7%; Pred. No. 3.2e-140;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60
Db 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60
QY 61 MLIGRYRDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLLCRLSLVYGE 120
Db 61 MLIGRYRDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLLCRLSLVYGE 120
QY 121 CTYATILHMTALSVERYLAICRPLRARVLTTRRRVALLAVMAVALLSAGPFLVGYE 180
Db 121 CTYATILHMTALSVERYLAICRPLRARVLTTRRRVALLAVMAVALLSAGPFLVGYE 180
QY 181 QDPGISVVGINGTARIASSPLASPPPLMSRAPPPSPSGPETAEAAALFSRECRPSA 240
Db 181 QDPGISVVGINGTARIASSPLASPPPLMSRAPPPSPSGPETAEAAALFSRECRPSA 240
QY 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGRQTVRL 300
Db 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGRQTVRL 300

RESULT 8
AAB62653
ID AAB62653 standard; Protein; 386 AA.

AC AAB62653;

DT 23-JUL-2001 (first entry)

DE Short form of motilin receptor, GPR-38B isoform.

KM zslg33; signal transduction; hormone; enzyme; neural development;
KM gastric contractility; nutrient uptake; digestive; pancreatic; human;
KM insulin-like growth factor-II; growth hormone; bone; gastrointestinal;
KM glucose; osteopathic; anorectic; vulnerability; immunomodulator; GHS-R;
KM G-protein coupled receptor; motilin receptor; GPR38; GPR-38B; isoform.

OS Homo sapiens.

PN WO200138355-A2.

PD 31-MAY-2001.

PF 22-NOV-2000; 2000WO-US32074.

PR 22-NOV-1999; 99US-0166765.

PA (ZYMO) ZYMOGENETICS INC.

PI Sheppard PO, Jaspers SR, Deisher TA, Bishop PD;

DR MPI; 2001-355879/37.

DN N-PSDB; AAF83684.

PT Forming reversible peptide receptor complex for purifying cell and
PT peptides, stimulating signal transduction and modulating hormone
PT secretion, involves contacting a receptor with zslg33 polypeptide
PS Disclosure; Page 106-109; 111pp; English.

CC The invention relates to a method of forming a reversible peptide-
CC receptor complex that involves providing an immobilized receptor, and
CC contacting the receptor with a zslg33 peptide (comprising residues 24-37
CC of AAB62649), where the receptor binds to the zslg33 peptide. The method
CC is useful for purifying cells, purifying a peptide, stimulating signal
CC transduction in a cell expressing a receptor. It is also useful for
CC modulating secretion of hormones, neural development and/or utilization,
CC gastric contractility, nutrient uptake, secretion of digestive and
CC pancreatic enzymes and hormones, secretion of insulin-like growth factor
CC -I, secretion of non-zslg33 proteins. It is useful for modulating growth

CC hormone secretion in a mammal having a disease associated with abnormal
CC levels of growth hormone, such as osteoporosis, bone repair, bone
CC remodeling, low osteoblast levels, cartilage repair and remodeling,
CC skeletal dysplasia, immune suppression, obesity, growth retardation,
CC protein catabolic responses after surgery, cachexia, protein loss,
CC dwarfism, wound healing and ovulation induction, treating a mammal having
CC a metabolic disorder requiring neurological feedback, such as satiety
CC regulation, glucose absorption and metabolism and neurotropy-associated
CC gastrointestinal disorders, and stimulating glucose-induced insulin
CC release in a mammal. The present sequence represents the short form of
CC motilin receptor, GPR-38B (one of the two isoforms of GPR38 which result
CC from alternative splicing). GPR38 has homology to the human G-protein
CC coupled receptor, GHS-R.

CC Sequence 386 AA;

Query Match 73.1%; Score 1575; DB 22; Length 386;
Best Local Similarity 99.7%; Pred. No. 3.2e-140;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60
Db 1 MGSPWNGSDGEGARBPMPALPPCDERRCSPPPLGALVPTAVCLCFVVGSGNVTV 60

QY 61 MLIGRYRDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLLCRLSLVYGE 120
Db 61 MLIGRYRDMRTTNLYGSMVSDLLILGLPDLRLMRSPWVGPLLCRLSLVYGE 120

QY 121 CTYATILHMTALSVERYLAICRPLRARVLTTRRRVALLAVMAVALLSAGPFLVGYE 180
Db 121 CTYATILHMTALSVERYLAICRPLRARVLTTRRRVALLAVMAVALLSAGPFLVGYE 180

QY 181 QDPGISVVGINGTARIASSPLASPPPLMSRAPPPSPSGPETAEAAALFSRECRPSA 240
Db 181 QDPGISVVGINGTARIASSPLASPPPLMSRAPPPSPSGPETAEAAALFSRECRPSA 240

QY 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGRQTVRL 300
Db 241 QLGALRWMLWTTAVFPLPCLSLTYGLIGRELMSRRPLRGPASGRERGRQTVRL 300

RESULT 9

AAB68476
ID AAB68476 standard; Protein; 271 AA.

AC AAB68476;

DT 23-JUL-2001 (first entry)

DE Amino acid sequence of a dog motilin receptor exon 1.

KM Motilin receptor; gastrointestinal disease; gastric motility disorder;
KM gastroparesis; irritable bowel syndrome; diarrhoea.

OS Canis sp.

PN WO200132710-A1.

PD 10-MAY-2001.

PF 25-OCT-2000; 2000WO-US29426.

PR 29-OCT-1999; 99US-0162264.

PA (MERI) MERCK & CO INC.

PI Tan C, McKee K;

DR MPI; 2001-343479/36.

DN N-PSDB; AAF85447.

PT Novel polypeptides related to dog and rabbit motilin receptor
PT polypeptide, comprising unique regions from dog and motilin receptor

PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans -
 XX
 PS Claim 4; Page 17; 42pp; English.
 CC The present sequence is encoded by exon 1 of a dog motilin receptor gene.
 CC The specification describes an unique sequence present in exon 1 of
 CC the motilin receptor, which is not present in human or Sphæroids
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 CC
 XX Sequence 271 AA;
 SQ
 Query Match 53.0%; Score 1143.5; DB 22; Length 271;
 Best Local Similarity 76.7%; Pred. No. 1.1e-99;
 Matches 230; Conservative 13; Mismatches 28; Indels 29; Gaps 4;
 QY 1 MGSPNNGSDGEGAREPPPALPCDERCSPPLGALVPTAVCLCFVVGSGNVMY 60
 Db 1 MGGPNSDGAEGAQ-----LPDERLCSPPLGALVPTAVCLCFVVGSGNVLTV 53
 QY 61 MLGRYDMRTTNNLYGSAVSDLLILGPDLYRLMSRPVGFPLCRSLTVYGG 120
 Db 54 LLIGRYDMRTTNNLYGSAVSDLLILGPDLYRLMSRPVGFPLCRSLTVYGG 113
 QY 121 CTYATLHMTLSVERYLACPLPARVLTTRRRALIAVMAVALSAGPFLVGYE 180
 Db 114 CTYATLHMTLSVERYLAVCPRLPARALLSRRAALIAVMAVALSAGPFLVGYE 173
 QY 181 ODPGISVVGNGTGRIRIASPPPLWLSRAPSPSPSGPETAAMALFSSECPSPA 240
 Db 174 QDAG---GPGNGSRRLRA-----PSPPGPE---AALFSRECPSPS 211
 QY 241 QLGALRWLWTTAVFPLPFLCLSLILYGLIGELMSRRPLRGPAASGRGRHQTKRV 300
 Db 212 QLGALRWLWTTAVFPLPFLCLCLVLYGRIGRELRRRGRPLRGPAASGRGRHQTKRV 271
 RESULT 10
 AAY54147
 ID AAY54147 standard; Protein, 363 AA.
 XX
 AC AAY54147;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE The puffer fish motilin receptor protein encoded by clone 75E7.
 XX
 XX Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;
 KW Applied form; MTL-R1A; MTL-R1B; gastric motility disorder;
 KW functional defect; neurological disorder; scleroderma; colonoscopy;
 KW paraneoplastic syndrome; radiation induced dysmotility; diabetes;
 KW infection; stress-related motility disorder; psychogenic disorder;
 KW gastroparesis; gastro-oesophageal reflux disease; constipation;
 KW chronic idiopathic pseudo obstruction; acute faecal impaction;
 KW postoperative ileus; gallstones; infantile colic; diarrhoea;
 KW irritable bowel syndrome; non-ulcer dyspepsia; non-cardiac chest pain;
 KW endoscopy; duodenal intubation.
 XX
 OS Spheroides nephelus.
 XX
 PN WO964436-A1.
 XX
 PD 16-DEC-1999.
 XX
 PF 08-JUN-1999; 99WO-US12773.
 XX
 PR 12-JUN-1998; 98US-0089098.

XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Feigener SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD,
 XX Pong S, Smith RG;
 DR WPI; 2000-105868/09.
 XX N-PSDB; AAZ45405.
 PT Novel receptor protein for screening compounds used in treating
 PT irritable bowel syndrome, constipation and other gastric conditions -
 XX
 PS Claim 7; Fig 8; 44pp; English.
 CC The present sequence represents a motilin receptor (clone 75E7) which
 CC is related to the human motilin receptor of the invention. The
 CC specification describes a G-protein coupled receptor, designated
 CC MTL-R1 (also GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see
 CC AAY54145) and MTL-R1B (see AAY54146). MTL-R1A is a functional seven
 CC transmembrane domain form, and MTL-R1B is a truncated five
 CC transmembrane domain. The MTL-R1 proteins are used to identify agonists
 CC and antagonists which can be used for treating gastric motility
 CC disorders, functional defects, disorders secondary to neurological
 CC disorders e.g. scleroderma, paraneoplastic syndromes radiation induced
 CC dysmotility, diabetes, infections, stress-related motility disorders,
 CC psychogenic disorders, gastroparesis, gastro-oesophageal reflux disease,
 CC constipation, chronic idiopathic pseudo obstruction, acute faecal
 CC impaction, postoperative ileus, gallstones, infantile colic, irritable
 CC bowel syndrome, non-ulcer dyspepsia, non-cardiac chest pain and
 CC diarrhoea. They can also be used in the preparation for colonoscopy,
 CC endoscopy and duodenal intubation.
 XX
 SQ Sequence 363 AA;
 Query Match 41.8%; Score 901; DB 21; Length 363;
 Best Local Similarity 50.3%; Pred. No. 1.3e-76;
 Matches 179; Conservative 56; Mismatches 77; Indels 44; Gaps 5;
 QY 31 SPPPLGALVVTAVVCLCFVVGSGNVMYVMTLIGRYDMRTTNNLYGSAVSDLLILG 90
 Db 31 SLFPASTLLEVTYICILIFVGVGTGNTTILIOYFDMKTNNLYSSMAVSDLVIFLC 90
 QY 91 LPEDLYRLMSRPVGFPLCRSLTVYGGCTYATLHMTLSVERYLACRPARVLT 150
 Db 91 LPEDLYRLMKYVPMFLGEANCRILYHIFECCTSAITLHMTLSVERYLACRPARVLT 150
 QY 151 TTRRRVALLAVMAVALSAGPFLVGYEQDPGISVVGNGTGRIRIASPPPLW 210
 Db 151 TTRRRVYIILALMCFALVSAAPTLFVGYEYD-----NET----- 185
 QY 211 SRAPSPSPSGPETAAMALFSRECPSPAQL--GALRWLWTTAVFPLPFLCLSLI 268
 Db 186 ----HPDYNTG-----QCKHTGYAISGQHLHMTWSTYFPCMLCLPLYG 229
 QY 269 LIGRELMSRRPLRGPAASGRGRHQTKRVLLVVLAFITICMLPFVHGRITVINTDSR 328
 Db 230 STGCKLMKSKNDQGCALABERSHQTVLIVVLAFFITICMLPFVHGRITVINTDSR 289
 QY 329 MWTFQYFNIVALLQFLYLGASTPIILYNLSKRYRAAFLLARKSRPGFHRSR 384
 Db 290 TMLLSQNFNMAWVLCYLSASINPVYVNLMSRYRAAKRLFLHQ-RKPAHRGQ 344
 RESULT 11
 AAB68479
 ID AAB68479 standard; Protein, 363 AA.
 XX
 AC AAB68479;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Amino acid sequence of a motilin receptor polypeptide.

KM Motilin receptor; gastrointestinal disease; gastric motility disorder;
 KM gastroparesis; irritable bowel syndrome; diarrhoea.
 OS Sphaeroides nephelus.
 XX WO200132710-A1.
 XX 10-MAY-2001.
 XX 25-OCT-2000; 2000WO-US29426.
 XX 29-OCT-1999; 99US-0162264.
 XX (MERI) MERCK & CO INC.
 XX Tan C, McKee K;
 XX WPI; 2001-343479/36.
 XX N-PSDB; AAF85450.
 DR Novel polypeptides related to dog and rabbit motilin receptor
 XX polypeptide, comprising unique regions from dog and motilin receptor
 PT amino acid sequence, useful for identifying compounds for treating
 PT diarrhoea in humans
 XX
 PS Disclosure; Page 33-34; 42pp; English.
 XX
 XX The present sequence represents a bacterial motilin receptor polypeptide.
 CC The specification describes an unique sequence present in exon 1 of
 CC the dog motilin receptor, which is not present in human or Sphaeroides
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.
 XX
 XX Sequence 363 AA;
 SQ
 Query Match 41.8%; Score 901; DB 22; Length 363;
 Best Local Similarity 50.3%; Pred. No. 1.3e-76;
 Matches 179; Conservative 56; Mismatches 77; Indels 44; Gaps 5;

AC AAY90666;
 XX 21-AUG-2000 (first entry)
 DT
 XX Human mutant G protein-coupled receptor GHSR (V262X).
 DE
 KM G protein-coupled receptor; GPCR; constitutively active;
 KM intracellular loop 3; transmembrane domain 6; drug screening;
 KM agonist; antagonist; mutant; mutein.
 XX Homo sapiens.
 OS Synthetic.
 OS
 PN WO200022129-A1.
 XX 20-APR-2000.
 XX 12-OCT-1999; 99WO-US23938.
 XX 13-OCT-1998; 98US-0170496.
 XX (AREN-) ARENA PHARM INC.
 XX Behan DP, Chalmers DT, Liaw CW;
 XX WPI; 2000-329165/28.
 DR N-PSDB; AAA30732.
 XX
 PT Non-endogenous constitutively activated human G protein-coupled
 PT receptor, useful for identifying agonists for use as pharmaceutical
 PT agents
 XX
 XX Example 2; Page 267-268; 341pp; English.
 PS
 CC The invention relates to constitutively active, non-endogenous versions
 CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
 CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743
 CC and AAA30775-A30779). The mutant proteins of the invention contain a
 CC mutation in a portion of the protein comprising intracellular loop 3
 CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
 CC is substituted for an endogenous residue in TM6 at a position 16 amino
 CC acids N-terminal of an endogenous proline in TM6 to form a sequence
 CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
 CC or Ala, and is preferably Lys. When the endogenous residue at this
 CC position is Lys, this residue is replaced by His, Arg or preferably Ala.
 CC The 15 amino acid stretch between the substituted amino acid and the Pro
 CC may be endogenous, non-endogenous, or a mixture of endogenous and
 CC non-endogenous residues. The constitutively active GPCRs are useful for
 CC identifying antagonists, agonists and partial agonists for use as
 CC pharmaceutical agents. The mutant proteins are also useful in research
 CC settings for elucidating the roles of the receptors in normal and
 CC diseased conditions. Antagonists for a particular GPCR are useful for
 CC treating diseases and disorders associated with that receptor. Because
 CC the novel mutant GPCRs are constitutively active, they can be used
 CC directly for screening of compounds without the need for endogenous
 CC ligands. Sequences AAY90643- AAY90677 and AAY90683-Y90687 the mutant
 CC human GPCRs of the invention.
 XX
 SQ Sequence 366 AA;
 Query Match 40.1%; Score 864.5; DB 21; Length 366;
 Best Local Similarity 44.7%; Pred. No. 3.7e-73;
 Matches 186; Conservative 59; Mismatches 106; Indels 65; Gaps 9;

DB 122 VSESCYIAVLTITLALSVETFAICPEPLRAKVVTKGVKLVIFVIAWVAFCSAGPIFVL 181
 QY 177 VGVEDDPGISVVPGLNGTARIASSPLASPPMLSRAPPSPPSGPTAEAAALFSRECR 236
 DB 182 VGVHEH-----NGT-----DP--W-----DINECR 199
 QY 237 PS--PAQCALKVMVMTTAYAFPLCLSLYLGLIGRELWSSRRPLRGPAASGERGRHR 294
 DB 200 PTEFAVRSGLLTVMVWSSIFPFLPVFCLTVLSLIGRKLIMRRRRGDVAVGASLRDQNHK 259
 QY 295 QTRKRLVLLVLAFLICMLPPIHVRRIIYNEDS--RMWTFQYFNVALQFLYSASIN 351
 DB 260 QTRKRLVAVVFAFLICMLPPIHVRRIIYFSKSFEPGSLERIAQISQYCNLVSEVLFYLSAIN 319
 QY 352 PLYVNLISKRYAAAFKLLARKSRPRGFRHSDTAGEVAGDTGDTVGTETSAN 407
 DB 320 PLYVIMSKRYAVAVFRLLGFEPPFSQKSLTKDESSR-----AMTESSIN 365

RESULT 13

AAy6293
 ID AAY6293 standard; Protein; 349 AA.

AC AAY6293;

DT 19-JUN-2000 (first entry)

DE A canine growth hormone secretagogue receptor (GHSR) protein.

XX Canine; growth hormone secretagogue receptor; GHSR; dog.

OS Canis sp.

PN WO200009538-A2.

PD 24-FEB-2000.

PF 06-AUG-1999; 99WO-US17915.

PR 10-AUG-1998; 98US-0095960.

PA (MERI) MERCK & CO INC.

PI Howard AD, Palyna OC, Smith RG, Tan CP;

DR WPI; 2000-224272/19.

DR N-PSDB; AAZ61492.

PT A novel canine growth hormone secretagogue receptor used to identify agonists and antagonists

PS Claim 6; Fig 2; 32pp; English.

XX The present sequence represents a canine growth hormone secretagogue receptor (GHSR). The GHSR cDNA sequence was isolated from a canine cDNA library constructed from spleen. The GHSR polynucleotide is a source of probes and primers, which can be used to isolate GHSR genes from other species. It is also used for recombinant expression of the polypeptide. The GHSR polypeptide is used in assays to identify agonists, antagonist and inhibitors. Such identified compounds can serve as leads for the development of veterinary pharmaceuticals that can be used to treat dogs having illnesses in which inappropriate activity of the canine GHSR proteins involved. The polypeptides may also be used to raise antibodies for diagnostic and therapeutic purposes.

XX Sequence 349 AA;

Query Match 39.9%; Score 859.5; DB 21; Length 349;

Best Local Similarity 46.2%; Pred. No. 1e-72;

Matches 166; Conservative 50; Mismatches 98; Indels 69; Gaps 8;

QY 14 AREPP-----WPALEPCDERRSCPPFLGALVPTAVVCLCFVVGSGNVVTVMLIGRYDM 69
 DB 6 AREPGSAGMDLP-----LFAPILAGVTATCVALLFAVGAGNLLTVRRFRRL 57
 QY 70 RTTNLYIGNAVSDLLILGLPFDLVRLMRSRPWFGLPCLSLYVGEQTYATLLHM 129
 DB 58 RTTNLYICSLACSDLLIFLCMPDLVRLMQRRPTFGDLCKLQFVSEGCYATVLI 117
 QY 130 TALVERTLACRLPRAVTVTRRVKLLIYVAVALLSAGPFLPLGVEDDPGISVVP 189
 DB 118 TALSVETFAICFPLRAKVLTKGVKLLAIWVAFCSAGPIFVLGVHEH----- 170
 QY 190 GINGTARIASSPLASPPMLSRAPPSPPSGPTAEAAALFSRECRPS--PAQCALAV 247
 DB 171 --NGT-----DPRD-----TRCRATEFVRSGLLTA 195
 QY 248 MLWTTAYFPLPCLSLYLGLIGRELWSSRRPLRGPAASGERGRHROTGRVLLVVLAF 307
 DB 196 MWVSVVFFFLPVFCLTVLYGLIGRKLIMRRRGDTPAGASLRDQNHK 255
 QY 308 IICMLPPIHVRRIIYNEDS--RMWTFQYFNVALQFLYSASINPLVNLISKRYA 364
 DB 256 IFCMLPPIHVRRIIYFSKSFEPGSLERIAQISQYCNLVSEVLFYLSAINPLVIMSKRYV 315
 QY 365 AAFKLLARKSRPRGFRHSDTAGEVAGDTGDTVGTETSAN 407
 DB 316 AVFKLLGFEPPFSQKSLTKDESSR-----AMTESSIN 348

RESULT 14

AAy90632
 ID AAY90632 standard; Protein; 366 AA.

AC AAY90632;

DT 21-AUG-2000 (first entry)

DE Human G protein-coupled receptor GHSR.

XX G protein-coupled receptor; GPCR; constitutively active;

KW intracellular loop 3; transmembrane domain 6; drug screening;

XX agonist; antagonist.

OS Homo sapiens.

PN WO200022129-A1.

PD 20-APR-2000.

PF 12-OCT-1999; 99WO-US23938.

PR 13-OCT-1998; 98US-0170496.

PA (AREN-) ARENA PHARM INC.

PI Behan DP, Chalmers DT, Liaw CW;

DR WPI; 2000-329165/28.

DR N-PSDB; AAA30643.

PT Non-endogenous constitutively activated human G protein-coupled receptors, useful for identifying agonists for use as pharmaceutical agents

PS Example 1; Page 167-169; 341pp; English.

XX The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-CC AAY90677 and AAY90683-990687), and to DNA encoding them (AAA30709-A30743-CC and AAA30775-A30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X, is substituted for an endogenous residue in IC3 at a position 16 amino

CC acids N-terminal of an endogenous proline in TM6 to form a sequence
 CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
 CC or Ala, and is preferably Lys. When the endogenous residue at this
 CC position is Lys, this residue is replaced by His, Arg or preferably Ala.
 CC The 15 amino acid stretch between the substituted amino acid and the Pro
 CC may be endogenous, non-endogenous, or a mixture of endogenous and
 CC non-endogenous residues. The constitutively active GPCRs are useful for
 CC identifying antagonists, agonists and partial agonists for use as
 CC pharmaceutical agents. The mutant proteins are also useful in research
 CC settings for elucidating the roles of the receptors in normal and
 CC diseased conditions. Antagonists for a particular GPCR are useful for
 CC treating diseases and disorders associated with that receptor. Because
 CC the novel mutant GPCRs are constitutively active, they can be used
 CC directly for screening of compounds without the need for endogenous
 CC ligands. The present sequence represents a human wild-type GPCR referred
 CC to in an exemplification of the invention.

XX Sequence 366 AA;

Query Match 39.8%; Score 857.5; DB 21; Length 366;
 Best Local Similarity 44.5%; Pred. No. 1.7e-72;
 Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

QY 5 WNGSDPEGA-----REPPPALPPCD---ERRCSPFGALVPYAVCLCFVGVSGN 56
 DB 2 WNAITBEEGPNLTADLDMDASPGNDSIGDELQFPAPLAGVATCAVAFVVGIAQN 61
 QY 57 VVTVMIGRYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLMRSPWVFGPLCRSLSY 116
 DB 62 LITMLVVSFRRLRTTNLYLSMAFSDLLIFLCMPDLVRLMQYRPMWFGDLCKLQFQ 121
 QY 117 VEGGCTYATLLHMTALSVRYLAICRPLARVLTERRRVALAVMAVALSAGPFLP 176
 DB 122 VSESTYATVLTITLALSVRYFAICRPLAKVVTGRKVLVFIWAAVAFSAGPIFVL 181
 QY 177 VGEODPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPPSGPETEAALFSRECR 236
 DB 182 VGVHE-----NGT-----DP--W-----DTNECR 199
 QY 237 PS--PAQLGALRYMLWTTAYFPLPFLCLSLYLGLIGRELMSSRRPLRGPAASGRGRHR 294
 DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCULTVLSLIGRKLRRRRGDVAVGASLRDQNHK 259
 QY 295 QTKRVLIVVLAFLICMLPHVGRITTYINTEDS---RMWYFSQYFNIALQFLYLSASIN 351
 DB 260 QTKMLAVVFAFLICMLPHVGRITLYFSKSPESGLEIAQISQYCNLVSFVFLSAIN 319
 QY 352 PILYNLSKKYRAAFAKLLARKSRPRGFRSRDAGEVADGTGDTVGYTETSAN 407
 DB 320 PILYNMSKKYRAVAFRLGFEPPSQKSLTKDESSR-----AWTESSIN 365

RESULT 15

AA70345 ID AAY70345 standard; Protein; 366 AA.

AC AAY70345;

DT 20-JUN-2000 (first entry)

DE Human G protein-coupled orphan receptor, GHSR.

XX G protein-coupled orphan receptor; GPCR, agonist; G protein; treatment;

KW GPCR fusion protein; inverse agonist; drug; GHSR; human.

XX Homo sapiens.

PN W0200006597-A2.

PD 10-FEB-2000.

PF 30-JUL-1999; 99WO-US17425.

PR 31-JUL-1998; 98US-0094879.
 PR 30-OCT-1998; 98US-0106300.
 PR 04-DEC-1998; 98US-0110906.
 PR 26-FEB-1999; 99US-0121651.

PA (AREN-) ARENA PHARM INC.

PI Behan DP, Chalmers DT, Liaw C, Lin I, Lowitz K, Chen R;

XX WPI, 2000-195260/17.

DR N-PSDB; AA251463.

XX Identification of a compound useful as a therapeutic agent, comprises

PT identifying a compound against constitutively activated G

PT protein-coupled orphan receptors

XX Example 2, Page 123; 123pp; English.

XX The patent discloses a method of identifying agonists and inverse or

CC partial agonists to the endogenous, constitutively activated

CC G protein-coupled orphan receptors (GPCRs), by contacting them with a

CC GPCR fusion protein comprising a GPCR and a G protein. Determining

CC expression of GPCRs in tissue samples can be used to identify related

CC diseases. Inverse agonists to these receptors can be used as drugs for

CC treating GPCR-related diseases. The present protein sequence is that of

CC human G protein-coupled orphan receptor, GHSR.

Sequence 366 AA;

Query Match 39.8%; Score 857.5; DB 21; Length 366;
 Best Local Similarity 44.5%; Pred. No. 1.7e-72;
 Matches 185; Conservative 59; Mismatches 107; Indels 65; Gaps 9;

QY 5 WNGSDPEGA-----REPPPALPPCD---ERRCSPFGALVPYAVCLCFVGVSGN 56
 DB 2 WNAITBEEGPNLTADLDMDASPGNDSIGDELQFPAPLAGVATCAVAFVVGIAQN 61
 QY 57 VVTVMIGRYRDMRTTNLYLGSMAVSDLLILGLPFDLYRLMRSPWVFGPLCRSLSY 116
 DB 62 LITMLVVSFRRLRTTNLYLSMAFSDLLIFLCMPDLVRLMQYRPMWFGDLCKLQFQ 121
 QY 117 VEGGCTYATLLHMTALSVRYLAICRPLARVLTERRRVALAVMAVALSAGPFLP 176
 DB 122 VSESTYATVLTITLALSVRYFAICRPLAKVVTGRKVLVFIWAAVAFSAGPIFVL 181
 QY 177 VGEODPGISVVPGLNGTARIASSPLASSPPLMLSRAPPSPPSGPETEAALFSRECR 236
 DB 182 VGVHE-----NGT-----DP--W-----DTNECR 199
 QY 237 PS--PAQLGALRYMLWTTAYFPLPFLCLSLYLGLIGRELMSSRRPLRGPAASGRGRHR 294
 DB 200 PTEFAVRSGLLTVMWVSSIFFLPVFCULTVLSLIGRKLRRRRGDVAVGASLRDQNHK 259
 QY 295 QTKRVLIVVLAFLICMLPHVGRITTYINTEDS---RMWYFSQYFNIALQFLYLSASIN 351
 DB 260 QTKMLAVVFAFLICMLPHVGRITLYFSKSPESGLEIAQISQYCNLVSFVFLSAIN 319
 QY 352 PILYNLSKKYRAAFAKLLARKSRPRGFRSRDAGEVADGTGDTVGYTETSAN 407
 DB 320 PILYNMSKKYRAVAFRLGFEPPSQKSLTKDESSR-----AWTESSIN 365

Search completed: March 16, 2003, 15:03:58
 Job time: 42 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2003, 00:44:59 ; Search time 3287 Seconds

(without alignments)
10051.352 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040
Sequence: 1 atgggcagccctggagcgg.....acgtgaagacgaltggagataa 2040

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estha.*
2: em_esthum.*
3: em_estlin.*
4: em_estmu.*
5: em_estrov.*
6: em_estrpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estlin.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.2	28.7	590	12	BF513101 UT-H-BW1-
2	255.8	12.5	425	17	AQ138681 HS_3073_A
3	226.8	11.1	500	12	BF603623 BF603623
4	142	7.0	504	13	A2451922 A2451922
5	127	6.2	855	17	BI757121 BI757121
6	127	6.2	1010	14	BM925480 BM925480

7	125.4	6.1	634	12	BG714306
8	125.4	6.1	1055	14	BM808829
9	125.2	6.1	540	14	BM737726
10	124.8	6.1	1027	13	BM543497
11	124	6.0	1144	13	BM546496
12	122.8	6.0	941	17	CNS036WH
13	119	5.8	1603	14	BM808823
14	117.8	5.8	624	9	AU244292
15	111.8	5.5	738	13	BI600170
16	111.2	5.5	928	13	BI758259
17	109.6	5.4	993	14	BI721315
18	109.2	5.4	448	12	BF323227
19	104.8	5.1	828	13	BI597845
20	102.6	5.0	980	13	BM543468
21	102.2	5.0	1189	11	AK005368
22	97	4.8	785	13	BI754749
23	97	4.8	843	17	CNS04SGU
24	96.8	4.7	670	17	AG046172
25	96.8	4.7	720	13	BI753905
26	96.6	4.7	555	12	BE751626
27	96.6	4.7	825	9	AU079556
28	94	4.6	1115	13	BM545572
29	93.6	4.6	959	17	CNS04GXP
30	92.2	4.5	1139	14	BQ056144
31	92.2	4.5	1139	14	BM922349
32	91	4.5	1798	17	AG171124
33	90.8	4.5	460	10	AW206493
34	89.6	4.4	624	10	BS587450
35	89.2	4.4	326	10	BE654108
36	89.2	4.4	462	10	AW138148
37	89.2	4.4	544	9	A1344017
38	89	4.4	741	13	BI224313
39	88.8	4.4	984	9	AL550664
40	88.4	4.3	961	17	CNS030RC
41	88.4	4.3	975	17	CNS039RV
42	88	4.3	866	9	AL553958
43	88	4.3	984	9	AL550996
44	87.8	4.3	1083	14	BM805382
45	87.6	4.3	1072	14	BQ055067

ALIGNMENTS

RESULT 1
LOCUS BF513101 590 bp mRNA linear EST 07-DEC-2000
DEFINITION UT-H-BW1-amn-e-07-0-UI.s1 NCI_CGAP_Sub7 Homo sapiens cDNA clone
IMAGE:3070549 3', mRNA sequence.

ACCESSION BF513101
VERSION BF513101.1 GI:11598280
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 590)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Oligo-dT track not found. Not 1 site shown in beginning of sequence is likely internal to the message. cDNA library preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution
Information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=No.

FEATURES
source 1.590
/organism="Homo sapiens"

Db 231 CATTTCCTGGGGGTGAGATGCTCCATGATGAGTTTCTCTAATTAATTGCTG-T 289

Qy 1611 ACTTGTATTCAGATGATGTTCTTGTGCGGGGATTAATTGCTTC 1661

Db 290 ACTTGTATTGAGATGTTCCCTTGTGCGGGTGGNGGTTTATGCTTCCC 340

RESULT 3
BF603623 500 bp mRNA linear EST 25-Apr-2001
LOCUS 269181 MARC 3HOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF603623
VERSION BF603623.1 GI:11701421
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 500)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Caasas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karayancheva,S., Liang,F., Quackenbush,J. and
Keeler,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mnscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 49 row: 1 column: 6
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1..500
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3HOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendonsus muscle, and fecal
longissimus muscle."

BASE COUNT 71 a 179 c 139 g 111 t

ORIGIN

Query Match 11.1%; Score 226.8; DB 12; Length 500;
Best Local Similarity 68.2%; Pred. No. 1.1e-39;
Matches 315; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 92 CGCCCTTCCCTGGGGGCGCTGCGCCGATACCGCTGTGCGCTGCTGCTG 151

Db 39 CCGCTTCCCGCGCGCTGCTGCGGGGCGTCACACCGCCTGCGGCTCTTGATG 98

Qy 152 TCGGGGTAGCGGCAACGATGATCGATGATCGGCGCTACCGGGAATCGGGA 211

Db 99 TGGGATGCGGGCACTGCTCACCATCTGTGTGTGCGCTTCCGGAGACTCGTA 158

Qy 212 CCACCACTTGTACTGTGGGACATGGCCGTGCCACCTACTCATCTGCTCGGGC 271

Db 159 CCACCACTTGTACTGTGCCATGGCCCTTCCGACTTACTCATCTTCTCTGCA 218

Qy 272 TGCCGTTGACCTGATGACCGCTCTGCGCGCTCGCGCCCTGGGTTGCGGCCGCTCT 331

Db 219 TGCCCTGACCTGCTGCGCTCTGCGCATTAACCGCCCTGGAACTTGGCGACTCTCT 278

Qy 332 GCCGCTGTCCCTTACGTGGGCGAGGGCTGCACCTACGCAACGCTGTGCATCATGCC 391

Db 279 GCAAACTCTCCAGTTTGTCAAGCGAGAGCTGCACCTACGCTACCGGTGCTACCATCACC 338

Qy 392 CGCTCAGCTGAGGCGCTACCTGCGCATCTGCGCCCGCTCCGCGCGCGCTTGTGTA 451

Db 339 CGCTGAGCTGAGGCTACTTCCCATCTGCTTCCGCTGCGGCGCAAGGTGTATCA 398

Qy 452 CCCGCGCGCGCTCGCGCGCTCATGCTGTGCTCTGCGCGCGCTGTCTTCTGCGC 511

Db 399 CCAAGGCGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 458

Qy 512 GTCCCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553

Db 459 GCGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 500

RESULT 4
AZ451922/c 504 bp DNA linear GSS 04-OCT-2000
LOCUS IM0251C06R Mouse 10kb plasmid UNGC1M library Mus musculus genomic
ACCESSION AZ451922
VERSION AZ451922.1 GI:10608187
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 504)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weis,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weis
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
94112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0251 row: 6 column: 06
Seq primer: CACACGGAACAGCTATGAC
Class: plasmid ends
High quality sequence file: 504.
Location/Qualifiers
1..504
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UNG1M0251C06"
/clone_lib="Mouse 10kb plasmid UNGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PMD42ny. Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (GI:473214|9d|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 156 a 110 c 107 g 131 t
ORIGIN

Query Match 7.0%; Score 142; DB 17; Length 504;
Best Local Similarity 75.6%; Pred. No. 6e-21;
Matches 189; Conservative 0; Mismatches 60; Indels 1; Gaps 1;

QY 1682 GTGCTGTCCTTATGTTGAGTGTGTGTTGTCGATTTATTAATTTGCTGGTTGCC 1740
DB 270 GTCTGTTTCCGAGCGATGTGTGAGTTTGGCATTCATAGTTGCTTGGTTGCTT 211
QY 1741 TTCCAGTTGGCGATCATTTATCAATAACGGAATTCGGGATGATGATCTTCT 1800
DB 210 TTCCAGTTGGCGAATCTTTTCATACACATAATTCGGGATGATGATCTTCT 151
QY 1801 CAGTACTTAACTCGTGTGCTGCACTTTTCTATCTGAGCGCATCTATCAACCAATC 1860
DB 150 CGATCTTTATATGATGATCTCTCAACTTCTATACCGAGTGCATCATCAAGCCATC 91
QY 1861 CTCTACACCTCATTTCAAGAGTACAGAGCGCGCTTTAAATCTGCTCTGCAAG 1920
DB 90 CTCCACACCTCATGTCATATGACAAAGCAGCTGTGAGAAATGCTTCTGCGAGA 31
QY 1921 AAGTCCAGGC 1930
DB 30 CAGTCCAGGC 21

RESULT 5

B1757121 855 bp mRNA linear EST 25-SEP-2001
LOCUS B1757121
DEFINITION 6030882F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200868 5',
mRNA sequence.

ACCESSION B1757121
VERSION B1757121
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 855)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLM1503 row: d column: 21
High quality sequence stop: 803.

FEATURES

source

Location/Qualifiers

1..855

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5200868"

/clone_lib="NIH_MGC_114"

/lab_host="DH10B"

/note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: Nct1;

BASE COUNT 103 a 300 c 299 g 153 t
ORIGIN

Query Match 5.2%; Score 127; DB 13; Length 855;
Best Local Similarity 55.2%; Pred. No. 1.3e-17;
Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

QY 27 CGGCCCCGAGGGGGGCGGAGCGCCGCTGCGCCGCTGCGCCCTTGCGAGAGCGCG 86
DB 60 CAGCAGCCCCGCGCCCGCCCGCGCGAGTCCAGCCCGGGGCTGAGCTGAGCGCCGCGCT 119
QY 87 CTGCTGCTTTTCCCTGCGGGCGCTGTGCGGCTGACCGCTGTGTGCTGCTGTT 146
DB 120 GGGCGTGAACCTGCGCTCTGCGGCGAAGTGTCTTTCACCGGCTCTTACGACTATCTG 179
QY 147 GCTGCTGGGGTGAAGCGGCAACGTG-----GTGACCGTATGATGAGGCGCTAACG 200
DB 180 GCGCTGGGCGCGGCGGCGAATGCGCTGTCCGTGACGTGTGTGAGGCGCGCGCG 229
QY 201 GACATGCGGACCAACCAACTTGTACTGCGGAGCATGCGCGCTGCGACTACTAT 260
DB 240 GCGCGCGGGGCGCGCGCGCGCACAGTGTGACGCTGCGGCGCTGCGCTGCTGCT 299
QY 261 CTGCTGCGGGCTGCGCGCTGTGACCTGTACCGCTCTGCGCGCTGCGG--CCCTGGCTGTT 317
DB 300 GCTGCTGCGCGCGCGCGGTGAGCTTCTACAGCTTGTGTCTTCACTACCTCTGGGCTT 359
QY 318 CGGCGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
DB 360 CGGCGACTGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
QY 378 GCTGACATGACCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 437
DB 420 GCTGAGCGTGGCAGGCGCTGAGCGCGCGAGCGCTGCTGAGCGCGCTGCTGCTGCTGCT 479
QY 438 CCGGCTCTTGTGACCGCGCGCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 497
DB 480 CCGGCGCTGCTGAGCTGCTGCTGCT 539
QY 498 GCTGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 550
DB 540 GCTGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592

RESULT 6

BW925480 1010 bp mRNA linear EST 12-MAR-2002
LOCUS BW925480
DEFINITION AGENCOURT.6625013 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5763267
5', mRNA sequence.

ACCESSION BW925480
VERSION BW925480.1 GI:19375859
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1010)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12814 row: n column: 04
 High quality sequence stop: 703.

FEATURES

source

Location/Qualifiers
 1..1010
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5763267"
 /clone_1ib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dr
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invtrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

BASE COUNT 149 a 348 c 337 g 176 t
 ORIGIN

Query Match 6.2%; Score 127; DB 14; Length 1010;
 Best Local Similarity 55.2%; Pred. No. 1.4e-17;
 Matches 294; Conservative 0; Mismatches 230; Indels 9; Gaps 2;

27 CGGCCCCAGAGGGGCGGAGAGCCCGCGCTGCGCCGCTTGGACGAGGCGCG 86
 Db CAGCAGCCCGGGGCGCCCGCGCCAGCTCCAAACCGGGGCTAGCGTGGACCGCGCT 141
 QY 87 CTGCTCGCCCTTTCCCTCGGGGCGCTGCTGCGCGCTGCTGCTGCTGCTGCT 146
 Db 142 GGGCGTGACACTCCCTCTGGCCAGAGTGTGTACCGCGCTTACGCACTATCTG 201
 QY 147 CGTGGCGGGGAGAGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 200
 Db 202 GGGCGTGAGCGCGGGGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 261
 QY 201 GGACATGCGGACCAACCAACTTGTACCTGGGAGCATGCGGCTTCCGCACTACTAT 260
 Db 262 GCGCGGGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 321
 QY 322 GCTGCTGGGCGTGGCGGCTGTGAGCTTCAAGCTTCTGTTGCTCACTACCTCGG 317
 Db 318 CGGCGCGGCTGCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
 Db 382 CGGCGGCTGCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 QY 378 GCTGACATGACCGCGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 437
 Db 442 GCTGAGGCTGACGCGCTGACGCGCTGACGCGCTGACGCGCTGACGCGCTGACG 501
 QY 438 CCGGCTGCTGCTGACCGCGCGCTGACGCGCTGACGCGCTGACGCGCTGACGCG 497
 Db 502 CCGGAGCTGCTGACCGCGCGCTGACGCGCTGACGCGCTGACGCGCTGACGCG 561
 QY 498 GCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 550
 Db 562 GCTGCGGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 614

RESULT 7
 BG714306 634 bp mRNA linear EST 08-MAY-2001
 LOCUS BG714306 602669812F1 NIH_MGC_96 Homo sapiens cDNA IMAGE:4792730 5',
 DEFINITION mRNA sequence.
 ACCESSION BG714306
 VERSION BG714306.1 GI:13993237

KEYWORDS

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 634)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10671 row: 9 column: 03

High quality sequence stop: 633.

Location/Qualifiers

1..634

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4792730"

/clone_1ib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (GCGAG

); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',

size-selected for average insert size 2.3 kb and

normalized to 10⁵. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

FEATURES

source

BASE COUNT 74 a 221 c 243 g 96 t
 ORIGIN

Query Match 6.1%; Score 125.4; DB 12; Length 634;
 Best Local Similarity 55.0%; Pred. No. 2.9e-17;
 Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

27 CGGCCCCAGAGGGGCGGAGAGCCCGCGCTGCGCCGCTTGGACGAGGCGCG 86
 Db 49 CAGCAGCCCGGGGCGCCCGCGCCAGCTCCAAACCGGGGCTAGCGTGGACCGCGCT 108
 QY 87 CTGCTCGCCCTTTCCCTCGGGGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCT 146
 Db 109 GGGCGTGACACTCCCTCTGGCCAGAGTGTGTACCGCGCTTACGCACTATCTG 168
 QY 147 CCGCTCGGGGAGAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 200
 Db 169 GGGCGTGAGCGCGGGGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 228
 QY 201 GGACATGCGGACCAACCAACTTGTACCTGGGAGCATGCGGCTTCCGCACTACTAT 260
 Db 229 GCGAGGCGGGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 288
 QY 261 CCGTGGGCGGCGGCTGACGCTGACGCGCTGACGCGCTGACGCGCTGACGCGCT 317
 Db 289 GCTGCTGCGCGGCGGCTGACGCTGACGCGCTGACGCGCTGACGCGCTGACGCG 348
 QY 318 CGGCGCGGCTGCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
 Db 349 CGGCGGCTGCTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408
 QY 378 GCTGACATGACCGCGCTGACGCGCTGACGCGCTGACGCGCTGACGCGCTGACG 437
 Db 409 GCTGAGCGTGGAGGCGCTGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 468

QY 438 CCGGCTTTGTGACCCGCGCGCGCTCCGCGCTCATCGCTGTCGTGGGCGGTGC 497
 Db 469 CCGGAGCTCTGTCAGCCAGCCGCGCGCGCTGGGCGCTCTCGTGGAGCGCTTC 528
 QY 498 GCTGCTCTCTGCGCGCTCTCTCTCTGTCGTGGGCGCTGAGAGACCCCG 550
 Db 529 GCTCGGCTCTGCGCGCTCTCTCTCTGTCGTGGGCGAGAGAGAGAGAGAGAG 581

RESULT 8
 BM808829 1055 bp mRNA linear EST 05-MAR-2002
 LOCUS 6617918 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5734336
 DEFINITION 5', mRNA sequence.
 ACCESSION BM808829
 VERSION BM808829.1 GI:19125652
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1055)
 NIH-MGC http://mgi.nci.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov

TISSUE Procurement: Invitrogen

CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov/
 Plate: L1M1273 row: h column: 17
 High quality sequence stop: 589.

FEATURES
 source
 1..1055
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5734336"
 /clone_id="NIH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: PCMV-SPORE; Site: 1: EcoRV
 (destroyed); Site: 2: NotI; RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."

BASE COUNT 160 a 376 c 336 g 183 t
 ORIGIN

Query Match 6.1%; Score 125.4; DB 14; Length 1055;
 Best Local Similarity 55.0%; Pred. No. 3, 1e-17;
 Matches 293; Conservative 0; Mismatches 231; Indels 9; Gaps 2;

QY 27 CCGGCCCCGAGGGGCGCGGAGCCCGCTGCGCGCTTCCGACGAGCGCG 86
 Db 53 CAGAGAGCCCGGCGCGCGCGCGAGCTCAACCCGCGCGCTGAGAGCGCGCT 112
 QY 87 CTGTGCGCTTTTCCCTGGGGGCGCGCTGAGCCGCTGTGCTGTGCTGT 146
 Db 113 GGGGGTGAACCTGCTCTGCGCGCAAGTGTGTCACCGGCTCTAGGACTATCTG 172
 QY 147 CGTGTGCGGGTGAAGCGGCAAGTGTGTCACCGGCTCTAGGACTATCTG 200
 Db 173 GGGCGTGGGCGCGCGCGGCAAGTGTGTCACCGGCTCTAGGACTATCTG 232
 QY 201 GAGCATGCGGACCAACCACTGTACTGCGGCGAGCATGCGCGCTGTCTACT 260

Db 233 GCGCGCGGGGCGCGCTGCGCGCACCACTGCTGACCTGCGCGCTGCGCTGCT 292
 QY 261 CCGTCTCGGAGTCCCGCTTGTGACCTGTACCGCTCTGCGCGCTCGCGG---CCCTGGGTGT 317
 Db 293 GCTGTGCTGCGCTGCGCGCTGTGAGCTTACAGCTTGTGTGTCTACATCCCTGGGTCTT 352
 QY 318 CCGGCGCGCTGCTGCGCGCTGTGCTTACGTGGGCGAGAGGCTGTACACCGGAGGCT 377
 Db 353 CCGGAGACTGTGCGCGCTGCGCGCTGTGCTTACGTGGGCGAGAGGCTGTGCGCGAGGCT 412
 QY 378 GCTGCAATGACCGCGCTGAGCGCGCTGAGCGCTACCTGAGCATGCGCGCGCTCGCGC 437
 Db 413 GCTGAGGTGCGAGCGCTGTGAGCGCGCGAGCGCTGCTGCTGCGCGCGCGCGCTGCG 472
 QY 438 CCGGCTCTGTCGACCGCGCGCGCGCGCTGCGCGCGCTGCTGCTGCTGCGCGCGCTGCG 497
 Db 473 CCGGAGCTGCTGTGACCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCGCGCGCTGCG 532
 QY 498 GCTGCTCTGCGCGCTGCTTCTTCTTCTGTCGTGGGCGGTGAGAGAGAGAGAGAG 550
 Db 533 GCTCGGCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585

RESULT 9
 BM737726 540 bp mRNA linear EST 01-MAR-2002
 LOCUS K-EST0000769 S1SNU5 Homo sapiens cDNA clone S1SNU5-16-A04 5', mRNA
 DEFINITION sequence.
 ACCESSION BM737726
 VERSION BM737726.1 GI:19059055
 KEYWORDS EST.

ORGANISM human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 540)
 Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

Kim, Y.S.
 JOURNAL 21C Frontier Korean EST Project 2001
 COMMENT Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.krribd.re.kr
 Plate: 16 row: A column: 04
 High quality sequence stop: 540.

FEATURES
 source
 1..540
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="S1SNU5-16-A04"
 /clone_id="S1SNU5"
 /sex="F"
 /tissue_type="Asciites"
 /cell_type="Lymphoblast-like"
 /cell_line="SNU-5"
 /lab_host="Top10F"
 /note="Organ: Stomach; Vector: PCMS; Site: 1: EcoRI;
 Site: 2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then dephosphorylated
 with tobacco acid pyrophosphatase (TAP). The dephosphorylated
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dT-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The

Journal of Management Inquiry 23(4) 391-407

CNMA Library Preparation: Life Technologies, Inc.
 CNMA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1M12726 row: e column: 02
 High quality sequence start: 27
 High quality sequence stop: 519.
 Location/Qualifiers

FEATURES

source

1..1144
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5729257"
 /clone_1ib="NH_MGC_124"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SORT6; Site 1: EcorV
 (destroyed); Site 2: NotI; RNA source male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcorV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."

BASE COUNT 188 a 384 c 371 g 200 t 1 others
 ORIGIN

Query Match 6.1%; Score 124; DB 13; Length 1144;
 Best Local Similarity 55.2%; Pred. No. 6.5e-17;

Matches 288; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

27 CCGCCCCGAGGGGGCGGGAGCCCGCTGCGCCCGCTTGGCAGCGAGCGCG 86
 67 CAGCAGCCCCCGGCGCTGCGGCCAGCTCCAGCCGGGCTGAGCGCCCGCT 126
 87 CTGCTGCGCCCTTCCCTGCGGGGCGCTGAGCGCGCTGAGCGCTGCTGCT 146
 127 GGGCGTGGACATCGCTCTGCGGCGAAGTCTGTTCAACCGCGCTTACGCACTATCTG 186
 147 CGTGTGCGGGGTGAGCGGCAACGTG-----GTGACCGTGTGCTGATCGGCGCTACCG 200
 187 GGGGCGGGGCGGGCGGCAATGCGTGTGCTGACAGTGTGCTGAAGCGCGGCGCG 246
 201 GGAACATGCGGACCAACCACTTGTACTGTGGAGAGATGGCGGTGTCCGACCTATCAT 260
 247 GCGCGCGGGGCGCTGCGCCACCACTGTCTAGCTGCGGCGCTGCGGGCGTGTCTCT 306
 261 CTGCTGCGGGGCTGCGCTGTGCACTGTACCGGCTC---TGGCGCTGCGGGCGCTGCTGTT 317
 307 GCTGTGCGGGTGGCTGAGCTTCTACAGCTTCTGTGTGTTCACTACCTCTGGGTCTT 366
 318 CCGGCGCGCTGCTGCGCGCTGTCTCTTACGTGGGCGAGGGCTGACCTTACGCGACGCT 377
 367 CCGGGAACCTGGGCTGCGCGGCTACTACTTCTGTGACAGAGTGTGCGCTACGCGACGCT 426
 378 GCTGACATGACCGCGCTCAAGCTGTAGCGCTACTGTGGCATTTGGCGCGCGCTCGCGG 437
 427 GCTGAGGTGTGCAAGCGCTGAGCGCGGAGCGCTGTGCTTACCGGTGTGCGCGCTGCTG 486
 438 CCGGCTTGTGTGCAACCGGCGCGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 497
 487 CCGGAGCTGTGTAAGCGCGACCGGAGCGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 546
 498 GCTGCTCTGCGCGGCTTCTTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 539
 547 GCTGTGCGCTGCGCTGTGCGCATGCGCGCTCATATGCGGGGAGA 588

RESULT 12
 CNS036WH/c 941 bp DNA linear GSS 15-MAY-2000
 LOCUS
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone

216F22 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 AL230570
 VERSION
 GSS; genome survey sequence.
 SOURCE
 ORGANISM

REFERENCE

1 (bases 1 to 941)
 Roest-Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Barnot, A., Fzimes, C., Wincker, P., Brothier, P., Quetier, F.,
 Saurin, W., and Weissbach, J.

TITLE

Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

REFERENCE

2 (bases 1 to 941)
 Roest-Crolius, H., Jallion, O., Dasilva, C., Fzimes, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

REFERENCE

3 (bases 1 to 941)
 Submitted (12-APR-2000)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the tetraodon nigroviridis
 genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

Location/Qualifiers

1..941
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="216F22"
 /clone_1ib="G"
 /note="Genoscope sequence ID : CNS0216DC11LP1-end : T7"

BASE COUNT 197 a 277 c 327 g 132 t 8 others
 ORIGIN

Query Match 6.0%; Score 122.8; DB 17; Length 941;
 Best Local Similarity 53.4%; Pred. No. 1.2e-16;
 Matches 318; Conservative 2; Mismatches 270; Indels 5; Gaps 3;

24 CAGAGCCCCGAGGGGGCGCGGAGCCCGGCGCGCGCGCGCTTGGCGAGCAGCG 83
 605 CAGAGCCACCGCGCGGAGACTTGAAGAGCGGCTGTGAGCGAGAGAGCGTACTTGGCCAG 546
 84 CCGGCTGCGCCCTTCCCTGCGGGGCGGTGCTGCGGCTGACCGCTGTGCTGTGCT 143
 545 GTACCGGGGCGCCCGCAGGTGCGCGCTTCTGCTGCGCGCTGTGCGCTACCTGGCAT 486
 144 GTTGTGTGTGCGGGGTGAGCGGCAACGTGTGACCGGTGATGTGCGGCTTACCGGGA 203
 485 CTTCCTGCGGGGCTGCGGCAACTCTCTGACCTGCGCGCTGATCTCTCGCTTCCGGCG 426
 204 CATGCGAACCAACCACTTGTACTGTGGGCAAGAGCGGTGTGCGACCTAC---TCAT 260
 425 GANCGCACGCGCCACCACTTACTTGTGTA-CTGTGGGCGGTGCGACCTGTGCTGTCT 367
 261 CCGTGTGCGGGCTGCGGCTTGTGACCTGTGACCGCTTGTGCGCGCTTGTGCTGTGCT 320
 366 GCTGTGCGGGAGTGCCTGTGAGCTGTGACACTGTGTGCGGAATTAACCTTCTGTGCTG 307
 321 GCGGCTGTCTGCGCGCTGCTTCTTACGTGTGCGGAGAGGCTGACCTTACGCGCTGCT 380
 306 GGAAGGGGGGCTGTACTTCCGACCTTCTGTTGAG-ACGCTGTGTTCCTCTCATCTT 248
 381 GCAATGACCGCGCTGAGGCTGAGCGCTTACCTGTGCGCATGTGCGCGCGCTTCCGCGCGCG 440


```
Db 270 CTTGCCCCCTGAGACTCTATAGATGTGACACACTACCCCTTCTGCTGAGCGTTGATG 329
Qy 330 CTGCGCCCTGCTCCCTTACGTGAGCGAGGCTGACCTAGCCAGCTGCTGACATGAC 389
Db 330 CTGCTATTTCGACACCTACTGTTTGTAGATGTCTGCTGCGCTCACTGCTCAACGTGAC 389
Qy 390 CCGGCTCAGGTGTGAGACCTGACCTGCGCATCTGCGCGCGCGCTGCGCGCGCTGCTG 449
Db 390 TGCCCTGAGGTGGAACGTATGTGCGGTGTGTCACCTCAGGCTCAGGTGCTCATGCT 449
Qy 450 CACCCGCGCGCGCGCTGCGCGCTGCTCATGCTGCTGCGCGCGCTGCTGCTGCTG 509
Db 450 GACGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCGCTGCGCGCTGCGCGCTGCTG 509
Qy 510 CCGTCCCTCTTCTGCTGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 568
Db 510 CCGTCCCAACACACGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 568
```

RESULT 2

```
US-10-083-168-82
; Sequence 82, Application US/10083168
; Publication No. US20030023069A1
; GENERAL INFORMATION:
; APPLICANT: Liaw, Chen W.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; APPLICANT: Maciejewski-Leinor, Dominique
; APPLICANT: Leonard, James N.
; APPLICANT: Ortuno, Daniel
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Endogenous And No. US20030023069A1-Endogenous, Constitutively Act
; FILE REFERENCE: AREN-0320
; CURRENT APPLICATION NUMBER: US/10/083.168
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: No. US20030023069A1el Sequence
US-10-083-168-82
```

Query Match 5.8%; Score 119; DB 9; Length 1212;
Best Local Similarity 54.5%; Pred. No. 3.5e-20;
Matches 261; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

```
Qy 93 GCGCTTCCCTGGGGCGGTGGCGGAGACCGCTGTCGCTGCTGCTGCTGCTGCTGCT 152
Db 90 GCGCCAGACAGACAGCTGTCATGCGCCATCTGTCACATACCTGTCGATCTTGTGCT 149
Qy 153 GCGGCTGAGCGGACAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 212
Db 150 GGGCGCTGTGGGCAATGGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 209
Qy 213 CACCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 269
Db 210 GCGTACCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 269
Qy 270 GCGTCCCTGAGACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 329
Db 270 CCGGCGCGCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 329
Qy 330 CTGCGCGCTGCTCCCTTACGTGAGCGAGGCTGACCTAGCCAGCTGCTGACATGAC 389
Db 330 CTGCTATTTCGACACCTACTGTTTGTAGATGTCTGCTGCGCTCACTGCTCAACGT 389
Qy 390 CCGGCTCAGGTGTGAGACCTGACCTGCGCATCTGCGCGCGCGCTGCGCGCGCTGCT 449
Db 390 TGCCCTGAGGTGGAACGTATGTGCGGTGTGTCACCTCAGGCTCAGGTGCTCATGCT 449
```

```
Db 390 TGCCCTGAGGTGGAACGTATGTGCGGTGTGTCACCTCAGGCTCAGGTGCTCATGCT 449
Qy 450 CACCCGCGCGCGCGCTGCGCGCTGCTCATGCTGCTGCGCGCGCTGCTGCTGCTG 509
Db 450 GACGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 509
Qy 510 CCGTCCCTCTTCTGCTGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 568
Db 510 CCGTCCCAACACACGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTG 568
```

RESULT 3

```
US-09-104-063-5
; Sequence 5, Application US/09104063
; Patent No. US20020168356A1
; GENERAL INFORMATION:
; APPLICANT: Lee, James
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PPA Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104.063
; FILING DATE: 24-June-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/701265
; FILING DATE: 22-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/664228
; FILING DATE: 06-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076093
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/810782
; FILING DATE: 19-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0706P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5530
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1679 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-104-063-5
```

Query Match 5.1%; Score 103.6; DB 9; Length 1679;
Best Local Similarity 48.7%; Pred. No. 3.1e-16;
Matches 348; Conservative 0; Mismatches 354; Indels 12; Gaps 2;

```
Qy 115 GTGCGGTGACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 174
Db 522 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
Qy 175 ACCGTGATGCTGATGCGGGGCTTACCGGAGATGCGGACACCACTTGTACTCTGGG 234
```

Db	582	GTGCTGGTGAATCTTGAGAGCGGCAACGGGACAAGACAGCAGTTCACAGGAGACCTTCCTGTTCC	641
Qy	235	AGCATGGCGGTGTCCGACCTTACTATCTCTGTCGGGCTGGCCGTTGCACCTGTACCGGCTTC	294
Db	642	CACCTGGCCGTTGGCCGACCTCTCTCTGTGTCTTCACTTTCGCCCTTTGACCGTG-----GCC	695
Qy	295	TGGCGCTCGCGGCGCCCTGGGAGTGTGGGCGCGCTGCTGCGCGCTGTCCCTTCAACGTGGGAC	354
Db	696	GAGGCGCTGTGGGGCTGGGTCTGGGAGCCTTCCTCTGTGAAACGTGATGTGCTTCGAC	755
Qy	335	GAGGCGCTGCACTTACGCCAGCGTGTGCACATGAACGCGGCTCAGCGTCAAGGCTTACTTG	414
Db	756	AAAGTCAACTTCTACTGTAGAGAGGCTTCTCTGGCCCTGCAATCCGCGTGAACGCTTACTTG	815
Qy	415	GCACATCTGCGCGCGCTCCGCGCGCGCGGTCTGTAGACCGGCGCGCGGTCCGCGCGCTC	474
Db	816	GCACATGTCCAGCGCGCTTCATGCTTACCGCAACGCGCGCTCTCTTCATTCACATCAGC	875
Qy	475	ATCGCTGTGCTTGTAGGCGCGTGGCGCGCTGCTTCGCGGTCCTTCTTGTTCGTGTGGGAC	534
Db	876	TGTGGGACATCTGGCTGGTGGGCTTCTCTTCCTTGCCCTTGCACAGATTCCTTTCGCAAA	935
Qy	535	GTCGAGCAGACACCCCGGCAATCTCGTATGTCCCGGCTCAATGACACCGCGGATGCGC	594
Db	936	GTCAGCCAAAGGCTATCAACAACATCTCCCTGCACAGTTGACCTTCTCCCAAGAACCAA	995
Qy	595	TCTCGCGCTTCGCGCTGCGCGCGCTCTCTGTGCTCTCGCGGCGCGC-----ACGCGC	648
Db	996	GCAGAAAGCATGTGCTGTTCACCTCCCATTTCTCTACATATGGCGGGATTTCTGTG	1055
Qy	649	TCCCGCGGTGGGGCCCGAGACCGCGGAGCGCGGCGCTGTTACGCCGAATGCGG	708
Db	1056	CCCATGTGCTGTGATGGGCTGGTGTACTAGTGGGGGTATGTGACAGTGTGGCCAGGCGCAG	1115
Qy	709	CCGAGCCCGCGGAGCTGTGGGCGCGCTCGGTATGTGTGGGTACCAACCGCTTACTTC	768
Db	1116	CGGCGCCCTCAGCGGCAAGAGCATGAGGTTGCCATCTGTGTCAAGCATTTCTTTC	1175
Qy	769	TTCCTGCCCTTCTGTGTGTCCTCAGCATCTCTACAGGCTCATCGGCGGAGGAGCTG	822
Db	1176	CTCTGTGTGTCACCTTACCATGTGTCATCTTCTGTGAACCTTGCGGAGGCTG	1229

RESULT 4
 US-09-823-114-18
 Sequence 18, Application US/09823114
 Patent No. US20020061554A1
 GENERAL INFORMATION:
 APPLICANT: EVANS, CHRISTOPHER J.
 TITLE OF INVENTION: OPTOID RECEPTOR GENES
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1888
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/823,114
 FILING DATE: 29-Mar-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/148,351
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:

```

1      NAME: MURASHIGE, KATE H.
2      REGISTRATION NUMBER: 29, 959
3      REFERENCE/DOCKET NUMBER: 22000-20526.22
4      TELECOMMUNICATION INFORMATION:
5      TELEPHONE: (202) 887-1500
6      TELEFAX: (202) 887-0763
7      TELEX: 90-4030 MRSNFOERSWSH
8
9      INFORMATION FOR SEQ ID NO: 18:
10     SEQUENCE CHARACTERISTICS:
11         LENGTH: 1805 base pairs
12         TYPE: nucleic acid
13         STRANDEDNESS: double
14         TOPOLOGY: linear
15
16     FEATURE:
17         NAME/KEY: CDS
18         LOCATION: 10..1119
19
20     SEQUENCE DESCRIPTION: SEQ ID NO: 18:
21     US-09-823-114-18

```

```

Query Match Similarity 5.0%; Score 102; DB 10; Length 1805;
Best Local Similarity 51.0%; Pred. No. 8,3e-16;
Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCACGCCCCCTGTGAACGGCAGCAGCGCCCGGAGGGGGCGCGGAGCGCGCGCTGAGCCGCG 63
DB 52 GGCACGCAACCTTCAGGGGCAACCTGTCCTCTGAGGCCCAACCAAGTCTGCTGCCCCCG 111
QY 64 CTGCGCGCTTGGCAGCAGCGCGCGCTGCTGCCCCCTTCCCTTGGGGGCGCTGATGCCGGT 123
DB 112 CATCTGCTGCTCAATGAGCCAGCAGCGCGCGCTTCTGTCGCCCTCGGGGCTCAAGGTCACATC 171
QY 124 ACCGCTGATGTCCTGTGCGCTGTTCTGTCCTGAGGGGTGAGGGGCAACGTGACCCGTGATG 183
DB 172 GTGGGGCTCTTACTCTGAGCGCTGTGTGTGTGGAGGAGCTCTTGGGAACTGCCCTGTGTCATG 231
QY 184 CTGATCGGGGCGCTACCGGAGCATGCGGACCAACCACTTGATCCTGGGACAGATAGCC 243
DB 232 GTGATCTTCAGGCAACCAAAATGAAGACAGCCACCAATATTACATCTTTAACCCTGGCC 291
QY 244 GTGTCCGACCTACTCATCTGCTCGGGCTGCGGTTGACCTGTGACCGCGCTTGGCGCTCG 303
DB 292 CTGGCGGACACTCTGTGTCTGTCTGAAGCTGCCCCCTT---CGAGGGACGGACATCTCTG 348
QY 304 CGGCGCTTGGGTGTGGGGCGCGCTGCTCTGCGCGCTGTCCCTCTACTGTGGGCGAGGGCTGC 353
DB 349 GCGCTTCTGCGCGTTTGGGAATGCGCTGTGCAAGACAGTATTGCCATTGACTACTACAAC 408
QY 364 ACGTACGGCAGCGTCTGTGCATGTACCGCGCTCAGCGTGTGAGCGCTGACTGCGCATCTGC 423
DB 409 ATGTTCACACAGCACTTCAACCTTAAGTGCATGAGTGTGATGCGCATATGACCATCTGC 468
QY 424 CGCGCGCTCCGCGCGCGCGTGTGTGATCACCGGCGCGCGGTCGCGCGCTCATCGCTGTG 483
DB 469 CACCCCATCTCGTGCGCTGTGACGTCCGACAGTCCAGCAAAAGCCAGGCTGTCAATGTGGCC 528
QY 484 CTCTGGGCGGTGCGCGCTGCTCTCTGCGCGGTGCCCTTCTTGTCTCTGG 529
DB 529 ATCTGGGCGCTGCGCTCTGTGTGTGCGGTGTCCGTTGCCATCATGG 574

RESULT 5
US-09-186A-9
US-09-905-186A-9
Sequence 9, Application US/0905186A
Publication No. US20030008289A1
GENERAL INFORMATION:
APPLICANT: Kreek, Mary Jeanne
APPLICANT: LaForge, Karl Steven
TITLE OF INVENTION: Alleles of the Human Orphanin
TITLE OF INVENTION: FQ/No. US20030008289A1:Orphanin Receptor Gene, Diagnostic Methods
FILE REFERENCE: 600-1-284N
CURRENT APPLICATION NUMBER: US/09/905,186A
CURRENT FILING DATE: 2001-10-19

```

PRIOR APPLICATION NUMBER: US 60/218,205
 PRIOR FILING DATE: 2000-07-14
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 9
 LENGTH: 1829
 TYPE: DNA
 ORGANISM: homo sapiens
 US-09-905-186A-9

Query Match
 Best Local Similarity 51.0%; Pred. No. 8.3e-16;
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCAGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 63
 DB 76 GGCAGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 135
 QY 64 CTGCGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 123
 DB 136 CATCTGCTCAATGCGAGCGACGCGCCCTTCTGAGCCCAACAGCTCTGCTGCGCGG 195
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
 DB 196 GTGGGGCTTACCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255
 QY 184 CTGATGCGGCGCTACCGGAGACGCGACCACTTGTACTTGGGAGCATGCGC 243
 DB 256 GTATCTCTGAGGACACCAAAATGAGACAGCAATATTATTTAATCTTAACTGCGC 315
 QY 244 GTGTCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
 DB 316 CTGCGCGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
 QY 304 CGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 DB 373 GCGCTTGGCGCTTGGGAATGCGCTGCAAGACAGTATTCATGACTGACTACAAAC 432
 QY 364 ACCTACGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 DB 433 ATGTTCACGACGACCTTACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
 QY 424 CGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
 DB 493 CACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
 QY 484 CTCTGGCGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
 DB 553 ATCTGGCGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

RESULT 6
 US-09-905-186A-10
 Sequence 10, Application US/09905186A
 Publication No. US20030008289A1
 GENERAL INFORMATION:
 APPLICANT: Kreek, Mary Jeanne
 APPLICANT: LaForge, Karl Steven
 TITLE OF INVENTION: Alleles of the Human Orphanin
 TITLE OF INVENTION: FO/No. US20030008289A1: Orphanin Receptor Gene, Diagnostic Methods
 TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon
 FILE REFERENCE: 600-1-284N
 CURRENT APPLICATION NUMBER: US/09/905,186A
 PRIOR FILING DATE: 2001-10-19
 PRIOR APPLICATION NUMBER: US 60/218,205
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 10
 LENGTH: 1829
 TYPE: DNA
 ORGANISM: homo sapiens
 US-09-905-186A-10

Query Match
 Best Local Similarity 51.0%; Pred. No. 8.3e-16;
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCAGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 63
 DB 76 GGCAGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 135
 QY 64 CTGCGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 123
 DB 136 CATCTGCTCAATGCGAGCGACGCGCCCTTCTGAGCCCAACAGCTCTGCTGCGCGG 195
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
 DB 196 GTGGGGCTTACCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255
 QY 184 CTGATGCGGCGCTACCGGAGACGCGACCACTTGTACTTGGGAGCATGCGC 243
 DB 256 GTATCTCTGAGGACACCAAAATGAGACAGCAATATTATTTAATCTTAACTGCGC 315
 QY 244 GTGTCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
 DB 316 CTGCGCGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
 QY 304 CGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 DB 373 GCGCTTGGCGCTTGGGAATGCGCTGCAAGACAGTATTCATGACTGACTACAAAC 432
 QY 364 ACCTACGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 DB 433 ATGTTCACGACGACCTTACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
 QY 424 CGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
 DB 493 CACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
 QY 484 CTCTGGCGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
 DB 553 ATCTGGCGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

RESULT 7
 US-09-905-186A-11
 Sequence 11, Application US/09905186A
 Publication No. US20030008289A1
 GENERAL INFORMATION:
 APPLICANT: Kreek, Mary Jeanne
 APPLICANT: LaForge, Karl Steven
 TITLE OF INVENTION: Alleles of the Human Orphanin
 TITLE OF INVENTION: FO/No. US20030008289A1: Orphanin Receptor Gene, Diagnostic Methods
 TITLE OF INVENTION: Alleles, and Methods of Treatment Based Thereon
 FILE REFERENCE: 600-1-284N
 CURRENT APPLICATION NUMBER: US/09/905,186A
 PRIOR FILING DATE: 2001-10-19
 PRIOR APPLICATION NUMBER: US 60/218,205
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 11
 LENGTH: 1829
 TYPE: DNA
 ORGANISM: homo sapiens
 US-09-905-186A-11

Query Match
 Best Local Similarity 51.0%; Pred. No. 8.3e-16;
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGCAGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 63
 DB 76 GGCAGCCCTTGGAAACGGCAGACGCGCCCGAGGGGCGCGGAGCCCGCTGCGCGG 135

QY 64 CTGCGCCCTTGAGAGAGCGCGCTGCTGCGCCCTTTCCCTTGAGGAGCGCTGCTGCGGTG 123
 Db 136 CATCTGCTGCTTAATGCCAGCCAGCGCGCTTCTGCGCCCTCGAGGCTCAAGGTCACCATC 195
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
 Db 196 GTGGGGCTCTACCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255
 QY 184 CTGATCGAGCGCTACCGGAGCATGCGGAGCACCACTTGTACTGTGAGCATGAGCC 243
 Db 256 GTCAATCTCAGGACACCAAAATGAAGACAGCACCAATATTATTAATCTTTAACTGGCC 315
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 303
 Db 316 CTGGCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
 QY 304 CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 Db 373 GGCCTTGGCGCTTGGGAAATGCGCTGTGCAAGACAGTATTCCTTACTACTACTCAAC 432
 QY 364 ACCTACCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 Db 433 ATGTTCACACACACCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTG 492
 QY 424 CGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
 Db 493 CACCCCATCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
 QY 484 CTCTGAGCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
 Db 553 ATCTGGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598

RESULT 8
 US-10-087-345A-22
 ; Sequence 22, Application US/10087345A
 ; Publication No. US20030045696A1
 ; GENERAL INFORMATION:

; APPLICANT: Owyang, Chung
 ; TITLE OF INVENTION: ORPHANIN FQ RECEPTOR NUCLEIC ACIDS
 ; FILE REFERENCE: UM-06962
 ; CURRENT APPLICATION NUMBER: US/10/087,345A
 ; CURRENT FILING DATE: 2002-03-01
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 2534
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-087-345A-22

Query Match 5.0%; Score 102; DB 9; Length 2534;
 Best Local Similarity 51.0%; Pred. No. 9.9e-16;
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGGAGCCCTTGAGAGAGCGCGCTGCTGCGCCCTTTCCCTTGAGGAGCGCGCTGCTGCGGTG 63
 Db 160 GGAAGCCACTTCAAGGCAACCTGCTCTCTGAGCCCAACCAAGTCTGTGCTGCTGCTG 219
 QY 64 CTGCGCGCTTGGAGAGAGCGCGCTGCTGCGCCCTTTCCCTTGAGGAGCGCGCTGCTGCGGTG 123
 Db 220 CATCTGCTGCTTAATGCCAGCCAGCGCGCTTCTGCGCCCTCGAGGCTCAAGTCACTATC 279
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
 Db 280 GTGGGGCTCTACCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
 QY 184 CTGATCGAGCGCTACCGGAGCATGCGGAGCACCACTTGTACTGTGAGCATGAGCC 243
 Db 340 GTATCTCTCAGGACACCAAAATGAAGACAGCACCAATATTATTAATCTTTAACTGGCC 399
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303

Db 400 CTGAGCCGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 456
 QY 304 CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 Db 457 GGCCTTGGCGCTTGGGAAATGCGCTGTGCAAGACAGTATTCCTTACTACTACTCAAC 516
 QY 364 ACCTACCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 Db 517 ATGTTCACACACCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTG 576
 QY 424 CGCGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
 Db 577 CACCCCATCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636
 QY 484 CTCTGAGCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
 Db 637 ATCTGGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 682

RESULT 9
 US-09-905-186A-1
 ; Sequence 1, Application US/09905186A
 ; Publication No. US20030008289A1
 ; GENERAL INFORMATION:

; APPLICANT: Kreek, Mary Jeanne
 ; APPLICANT: LaForge, Karl Steven
 ; TITLE OF INVENTION: Alleles of the Human Orphanin
 ; TITLE OF INVENTION: FQ/No. US20030008289A1 Orphanin Receptor Gene, Diagnostic Methods
 ; FILE REFERENCE: 600-1-284N
 ; CURRENT APPLICATION NUMBER: US/09/905,186A
 ; PRIOR FILING DATE: 2001-10-19
 ; PRIOR APPLICATION NUMBER: US 60/218,205
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2602
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-09-905-186A-1

Query Match 5.0%; Score 102; DB 9; Length 2602;
 Best Local Similarity 51.0%; Pred. No. 1e-15;
 Matches 268; Conservative 0; Mismatches 255; Indels 3; Gaps 1;

QY 4 GGGAGCCCTTGAGAGAGCGCGCTGCTGCGCCCTTTCCCTTGAGGAGCGCGCTGCTGCGGTG 63
 Db 731 GGAAGCCACTTCAAGGCAACCTGCTCTCTGAGCCCAACCAAGTCTGTGCTGCTGCTG 790
 QY 64 CTGCGCGCTTGGAGAGAGCGCGCTGCTGCGCCCTTTCCCTTGAGGAGCGCGCTGCTGCGGTG 123
 Db 791 CATCTGCTGCTTAATGCCAGCCAGCGCGCTTCTGCGCCCTCGAGGCTCAAGTCACTATC 850
 QY 124 ACCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
 Db 851 GTGGGGCTCTACCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910
 QY 184 CTGATCGAGCGCTACCGGAGCATGCGGAGCACCACTTGTACTGTGAGCATGAGCC 243
 Db 911 GTCAATCTCAGGACACCAAAATGAAGACAGCACCAATATTATTAATCTTTAACTGGCC 970
 QY 244 GTGTCCGACCTACTCATCTGCTGCGGCGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
 Db 971 CTGGCGGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027
 QY 304 CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 Db 1028 GGCCTTGGCGCTTGGGAAATGCGCTGTGCAAGACAGTATTCCTTACTACTACTCAAC 1087
 QY 364 ACCTACCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 Db 1088 ATGTTCACACACCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTG 1147

Query Match	4.6%	Score 93.8	DB 10	Length 1258
Best Local Similarity	51.1%	Pred. No. 8e-14		
Matches 246, Conservative	0	Mismatches 232	Indels 3	Gaps 1

```

RESULT 13
US-09-967-768A-296
; Sequence 296, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Self
; FILE REFERENCE: 689230-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28

```

Query Match	4.6%	Score 93.6	DB 10	Length 1427
Best Local Similarity	52.8%	Pred. No. 9.6e-14		
Matches 226	Conservative 0	Mismatches 199	Indels 3	Gaps 1

RESULT 14
US-10-112-599A-3

Query Match	4.5%	Score	92.6	DB	9	Length	1773
Best Local Similarity	52.4%	Pred. NO.	1.9e-13				
Matches	228	Conservative	0	Mismatches	204	Indels	3
						Gaps	1

OY 95 CATTTCCTGGGAGCGTGTGCCTGATTCGTGCG 154
||| ||| |
Db 361 CCTCGCTCCCTCGCCTTGGAATCGCATCAACGCCGCTACTCGGCCGTGTGCGCCTGG 420

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2003, 01:14:09 ; Search time 84 Seconds

(without alignments)
7447.864 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040

Sequence: 1 atgggacgacccctggaacg.....acgtgaagacgatggatata 2040

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	13.9	283	4	US-08-993-088A-4
2	283	13.9	283	4	US-08-993-424B-4
3	234.8	11.5	1063	4	US-09-077-675A-1
4	230	11.3	1029	4	US-09-077-675A-4
5	228.8	11.2	1122	4	US-09-077-675A-9
6	228.4	11.2	1092	4	US-09-077-675A-15
7	228.4	11.2	3129	4	US-09-077-675A-14
8	227.8	11.2	1088	4	US-09-077-675A-6
9	153	7.5	836	4	US-09-077-675A-11
10	134.8	6.6	1248	4	US-09-545-944-1
11	127	6.2	1575	3	US-08-858-876A-1
12	127	6.2	1575	3	US-09-472-880-1
13	125.4	6.1	1342	3	US-08-833-399-1
14	125.4	6.1	1342	4	US-09-372-498-1
15	122.2	6.0	1529	3	US-08-858-876A-3
16	122.2	6.0	1529	4	US-09-472-880-3
17	120	5.9	1535	4	US-09-668-680-12
18	115.4	5.7	1176	4	US-09-200-090-3
19	115.4	5.7	1233	4	US-09-200-090-1
20	109.8	5.4	1164	4	US-08-993-088A-6
21	109.8	5.4	1164	4	US-08-993-424B-6
22	106.8	5.2	1116	4	US-08-993-088A-18
23	106.8	5.2	1116	4	US-08-993-088A-19
24	106.8	5.2	1116	4	US-08-993-424B-18
25	106.8	5.2	1119	2	US-08-626-685A-7
26	106.8	5.2	1690	4	US-08-665-034A-1
27	106.6	5.2	1610	1	US-08-056-051-5

28	106.6	5.2	1610	1	US-07-928-611-21	Sequence 21, Appl
29	106.6	5.2	1610	2	US-08-487-811A-21	Sequence 21, Appl
30	106.6	5.2	1610	4	US-09-060-694-21	Sequence 21, Appl
31	106.6	5.2	1610	4	US-09-378-074-21	Sequence 21, Appl
32	106.6	5.2	1610	5	PCT-US93-07370-21	Sequence 21, Appl
33	103.6	5.1	1679	1	US-08-202-056-6	Sequence 6, Appl
34	103.6	5.1	1679	1	US-08-076-093A-5	Sequence 5, Appl
35	103.6	5.1	1679	1	US-08-701-265-5	Sequence 5, Appl
36	103.6	5.1	1679	2	US-08-284-586-5	Sequence 5, Appl
37	103.6	5.1	1679	2	US-08-805-478-5	Sequence 5, Appl
38	103.6	5.1	1679	2	US-08-802-627A-5	Sequence 5, Appl
39	103.6	5.1	1679	2	US-08-801-238-5	Sequence 5, Appl
40	103.6	5.1	1679	2	US-08-801-228-5	Sequence 5, Appl
41	103.6	5.1	1679	3	US-09-104-296-5	Sequence 5, Appl
42	103.6	5.1	1679	5	PCT-US94-06380-3	Sequence 3, Appl
43	103.6	5.1	2818	3	US-08-982-493-7	Sequence 7, Appl
44	103.6	5.1	2818	4	US-08-628-655-1	Sequence 1, Appl
45	102	5.0	1805	4	US-08-405-271A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-993-088A-4
Sequence 4, Application US/08993088A
Patent No. 6287855
GENERAL INFORMATION:
APPLICANT: Tan, Carina
TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,088A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,851
FILING DATE: 27-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 19846
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1958
TELEFAX: 732-594-4720
TEXT:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 283 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
FEATURE:
NAME/KEY: Other
LOCATION: 1...283
OTHER INFORMATION: cDNA probe
US-08-993-088A-4

Query Match 13.9%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGACCAACCACTTGTACTGGGAGAGATGAGCGGTGTCCGACTTACTATCTCTGC 265
DB 1 TGGGAGACCAACCACTTGTACTGGGAGAGATGAGCGGTGTCCGACTTACTATCTCTGC 60
QY 266 TGGGAGTGGCGTGTGACCTGTACCGCTCTGCGGCTCGGCGGCTGGGGTGTGGGGCGGC 325
DB 61 TGGGAGTGGCGTGTGACCTGTACCGCTCTGCGGCTCGGCGGCTGGGGTGTGGGGCGGC 120
QY 326 TGGTGTGCGCGCTGTCCCTCTACGTGGAGGAGGCTGACCTTACGCAAGCTGTGACACA 385
DB 121 TGGTGTGCGCGCTGTCCCTCTACGTGGAGGAGGCTGACCTTACGCAAGCTGTGACACA 180
QY 386 TGACCGCGCTAGCGGTGAGCGCTACTGCGCATCTGCGCGGCTCGGCGGCTGTCT 445
DB 181 TGACCGCGCTAGCGGTGAGCGCTACTGCGCATCTGCGCGGCTCGGCGGCTGTCT 240
QY 446 TGGTCAACCGCGCGCGGCTCGGCGGCTCATGCTGTGCTCTG 488
DB 241 TGGTCAACCGCGCGCGGCTCGGCGGCTCATGCTGTGCTCTG 283

RESULT 2

US-08-993-424B-4
Sequence 4, Application US/08993424B
Patent No. 6337206
GENERAL INFORMATION:
APPLICANT: Tan, Carina
APPLICANT: Kolakowski, Lee F., Jr.
TITLE OF INVENTION: MOUSE GLANIN RECEPTOR GALR2 AND
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,424B
FILING DATE: 18-DEC-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,851
FILING DATE: 27-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,119
REFERENCE/DOCKET NUMBER: 19846NP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1958
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 283 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
FEATURE:
NAME/KEY: Other
LOCATION: 1...283
OTHER INFORMATION: cDNA probe

US-08-993-424B-4
Query Match 13.9%; Score 283; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 3.6e-54;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TGGGAGCAACCACTTGTACTGGGAGAGATGAGCGGTGTCCGACTTACTATCTCTGC 265
DB 1 TGGGAGCAACCACTTGTACTGGGAGAGATGAGCGGTGTCCGACTTACTATCTCTGC 60
QY 266 TGGGAGTGGCGTGTGACCTGTACCGCTCTGCGGCTCGGCGGCTGGGGTGTGGGGCGGC 325
DB 61 TGGGAGTGGCGTGTGACCTGTACCGCTCTGCGGCTCGGCGGCTGGGGTGTGGGGCGGC 120
QY 326 TGGTGTGCGCGCTGTCCCTCTACGTGGAGGAGGCTGACCTTACGCAAGCTGTGACACA 385
DB 121 TGGTGTGCGCGCTGTCCCTCTACGTGGAGGAGGCTGACCTTACGCAAGCTGTGACACA 180
QY 386 TGACCGCGCTAGCGGTGAGCGCTACTGCGCATCTGCGCGGCTCGGCGGCTGTCT 445
DB 181 TGACCGCGCTAGCGGTGAGCGCTACTGCGCATCTGCGCGGCTCGGCGGCTGTCT 240
QY 446 TGGTCAACCGCGCGCGGCTCGGCGGCTCATGCTGTGCTCTG 488
DB 241 TGGTCAACCGCGCGCGGCTCGGCGGCTCATGCTGTGCTCTG 283

RESULT 3

US-09-077-675A-1
Sequence 1, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Peng, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1063 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

QY 212 TCCGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 331
Db 284 TCCGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 343
QY 332 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 391
Db 344 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 403
QY 392 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 451
Db 404 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 463
QY 452 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 511
Db 464 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 523
QY 512 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 553
Db 524 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 565

RESULT 7

US-09-077-675A-14
; Sequence 14, Application US/09077675A
; Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/077,675A

FILING DATE: 3-JUN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 3129 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-077-675A-14

Query Match 11.2%; Score 228.4; DB 4; Length 3129;
Best Local Similarity 68.4%; Pred. No. 1e-41;
Matches 316; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 92 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 151
Db 104 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 163
QY 152 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 211
Db 164 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 223
QY 212 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 271
Db 224 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 283
QY 272 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 331
Db 284 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 343
QY 332 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 391
Db 344 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 403
QY 392 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 451
Db 404 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 463
QY 452 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 511
Db 464 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 523
QY 512 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 553
Db 524 GCGGCTTGACCTGTACCGCCTCTTGCGCCTTGCGGCTTTCGAGCCGCTGCTCT 565

RESULT 8

US-09-077-675A-6
; Sequence 6, Application US/09077675A
; Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway

STATE: NJ

COUNTRY: USA

ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/077,675A

FILING DATE: 3-JUN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:

NAME: Cocuzzo, Anna L.

REGISTRATION NUMBER: 42,452

REFERENCE/DOCKET NUMBER: 19590P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 732-594-1273

TELEFAX: 732-594-4720

TELEX:

```

RESULT 10
US-09-545-944-1
: Sequence 1, Application US/09545944
: Patent No. 6461836
: GENERAL INFORMATION:
: APPLICANT: AMES, ROBERT
: APPLICANT: ELSHOUBAGY, NABIL
: APPLICANT: MICHALOVICH, DAVID
: APPLICANT: SARAU, HENRY
: APPLICANT: SHABON, USMAN
: APPLICANT: VAMTER, LISA
: TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
: TITLE OF INVENTION: (AXOR34) AND SCREENING METHODS THEREOF
: FILE REFERENCE: GP/0657-1
: CURRENT APPLICATION NUMBER: US/09/545, 944
: CURRENT FILING DATE: 2000-04-10
: PRIOR APPLICATION NUMBER: US 09/435,384
: PRIOR FILING DATE: 1999-11-05
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PasteSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1248
: TYPE: DNA

```

Query	DB	Score	Length	DB	Score	Length	DB	Score	Length
1	27	6.24	1575	2	6.24	1575	3	6.24	1575
2	45	55.24	1575	4	55.24	1575	5	55.24	1575
3	87	0	1575	6	0	1575	7	0	1575
4	105	0	1575	8	0	1575	9	0	1575
5	147	0	1575	10	0	1575	11	0	1575
6	165	0	1575	12	0	1575	13	0	1575
7	201	0	1575	14	0	1575	15	0	1575
8	225	0	1575	16	0	1575	17	0	1575
9	261	0	1575	18	0	1575	19	0	1575
10	285	0	1575	20	0	1575	21	0	1575
11	318	0	1575	22	0	1575	23	0	1575
12	345	0	1575	24	0	1575	25	0	1575
13	378	0	1575	26	0	1575	27	0	1575
14	405	0	1575	28	0	1575	29	0	1575
15	438	0	1575	30	0	1575	31	0	1575
16	465	0	1575	32	0	1575	33	0	1575
17	498	0	1575	34	0	1575	35	0	1575
18	525	0	1575	36	0	1575	37	0	1575

Db 238 GCGCCTGCGCTACCAAGTGTCTAGCCTGGCGCTCTGACCCCTGTGCTACTGCTGTCTAG 297
QY 270 GCTGCGCTTCTGACCTGTACCGCCTCTGGCGCTGGCGG--CCCTGGGGTGTTCGGGGCGCT 326
Db 298 CATGCCCATGAGAGCTCTACAACTTCGTGTGTCCACTTACCCATGGGTCTTCGGCGATCT 357
QY 327 GCTTGGCGCCTGTGCTCCCTCTAGCTGGGCGAGGGCTGCACCTTACGCCACAGCTGTGCACAT 386
Db 358 GGGCTGCGCTGCTATTACTTCTGCGCGGAGCTGTGCGCTAGCGCCACAGTGTGAGCGT 417
QY 387 GACCGGCTCAGCGTGAAGGCTTACCTGAGCCATCTGCCGCCCTCCGCCGCCGCTT 446
Db 418 TGCAGGCTTAAGGCGAGAGGCTGCTGCGCGGTGTGTCCAGCCGCTGCGCGCGCCGCT 477
QY 447 GGTCAACCGCGCGCGCGCTCCGCGCGCTCATGCTGTGCTCTGGGCGGTGGGCTGTCTC 506
Db 478 TCTCACCCCGCGCGCGCACCCGCGCGCTGTGTCTGTGTCTGTGTGTGTGTGTGTGTGT 537
QY 507 TGCCTGCTCCCTTCTTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 545
Db 538 TGCCCTGCCCATGGCGGTTATCATGGGACAGAAAGCACGA 576

Search completed: March 16, 2003, 10:51:09
Job time : 88 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 08:50:18 ; Search time 480 seconds
(without alignments)
9570.993 Million cell updates/sec

Title: US-09-876-252-129
Dafact score: 2040

Sequence: 1 atgggcagccctggaagg.....acgtgaagacgatgggataa 2040

Scoring table: IDENTITY_NUC

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

1. N_ /SID52/gcgdata/genseq/genseqn-emb1/NA1960.DAT.*
2. /SID52/gcgdata/genseq/genseqn-emb1/NA1961.DAT.*
3. /SID52/gcgdata/genseq/genseqn-emb1/NA1962.DAT.*
4. /SID52/gcgdata/genseq/genseqn-emb1/NA1963.DAT.*
5. /SID52/gcgdata/genseq/genseqn-emb1/NA1964.DAT.*
6. /SID52/gcgdata/genseq/genseqn-emb1/NA1965.DAT.*
7. /SID52/gcgdata/genseq/genseqn-emb1/NA1966.DAT.*
8. /SID52/gcgdata/genseq/genseqn-emb1/NA1967.DAT.*
9. /SID52/gcgdata/genseq/genseqn-emb1/NA1968.DAT.*
10. /SID52/gcgdata/genseq/genseqn-emb1/NA1969.DAT.*
11. /SID52/gcgdata/genseq/genseqn-emb1/NA1970.DAT.*
12. /SID52/gcgdata/genseq/genseqn-emb1/NA1971.DAT.*
13. /SID52/gcgdata/genseq/genseqn-emb1/NA1972.DAT.*
14. /SID52/gcgdata/genseq/genseqn-emb1/NA1973.DAT.*
15. /SID52/gcgdata/genseq/genseqn-emb1/NA1974.DAT.*
16. /SID52/gcgdata/genseq/genseqn-emb1/NA1975.DAT.*
17. /SID52/gcgdata/genseq/genseqn-emb1/NA1966.DAT.*
18. /SID52/gcgdata/genseq/genseqn-emb1/NA1977.DAT.*
19. /SID52/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20. /SID52/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21. /SID52/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22. /SID52/gcgdata/genseq/genseqn-emb1/NA2001.DAT.*
23. /SID52/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
24. /SID52/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2040	100.0	2040	21	AAA46116	Human G protein coupled
2	2022.2	99.1	3066	21	AAZ45402	Genomic sequence of
3	1047.2	51.3	1161	22	AAF83684	Short form of molet
4	1047.2	51.3	1390	21	AAZ45404	CDNA encoding the
5	896.2	43.9	1239	21	AAZ45403	CDNA encoding the
6	896.2	43.9	1239	21	AAF83683	long form of molet
7	896.2	43.9	1239	22	AAF85449	Nucleotide sequence
8	896.2	43.9	1506	22	AAI65089	Human GPR38 variant
9	669.8	32.8	1203	22	AAF85448	Nucleotide sequence

10	541.2	26.6	813	22	AAFS94447	Nucleotide sequence	
C	11	501.2	24.6	1179	24	ABQ47146	Oligonucleotide for
C	13	501.2	24.6	1179	24	ABQ47147	Oligonucleotide for
C	14	437.2	21.4	1179	24	ABQ47148	Oligonucleotide for
C	15	283	13.9	283	19	AAV44930	Oligonucleotide for
16	283	13.9	283	19	AAV34651	Galanin receptor G	
17	283	13.9	283	19	AAV28290	Galanin receptor G	
18	283	13.9	283	24	ABK14060	Galanin receptor G	
19	248.2	12.2	1050	21	AAZ61492	cDNA encoding cana	
20	234.8	11.5	1063	18	AAI68662	Pig growth hormone	
21	234.8	11.5	1063	18	AAI68754	Swine growth hormone	
22	233.2	11.4	1029	18	AAI68663	Pig growth hormone	
23	233.2	11.4	1029	18	AAI68755	Swine growth hormone	
24	231.6	11.4	1095	21	AAZ45993	cDNA encoding the	
25	231.6	11.4	4009	21	AAZ45967	DNA encoding the w	
26	228.8	11.2	1122	18	AAI68665	Human growth hormo	
27	228.4	11.2	1092	22	AAH27800	Human growth hormo	
28	228.4	11.2	3129	18	AAI68667	Rat growth hormone	
29	228.4	11.2	3129	18	AAI68759	Rat growth hormone	
30	227.8	11.2	1088	18	AAI68664	Human growth hormo	
31	227.8	11.2	1088	18	AAI68756	Human growth hormo	
32	227.8	11.2	1101	21	AAA30643	Human G-protein-co	
33	227.8	11.2	1101	21	AAA30732	DNA encoding human	
34	227.8	11.2	1101	21	AAZ51463	Human G-protein-co	
35	227.8	11.2	1101	21	AAAF83680	Human G-protein co	
36	227.8	11.2	1122	24	AAAD30395	Human growth hormo	
37	227.2	11.1	1122	18	AAI68757	Human growth hormo	
38	226.8	11.1	1092	18	AAI68760	Rat growth hormone	
39	225.2	11.0	1092	21	AAZ44405	DNA encoding the p	
40	225.2	11.0	1092	22	AAI685450	Nucleotide sequenc	
41	184.6	9.0	1098	22	AAAF83681	Human G-protein co	
42	153	7.5	836	18	AAI68666	Human growth hormo	
43	153	7.5	836	18	AAI68758	Human growth hormo	
44	134.8	6.6	729	22	AAI68107	Nucleotide sequenc	
C	45	134.8	6.6	801	22	AAH50977	Human nGPR15 codi

ALIGNMENTS

	RESULT	1
AC	AAA46116	
ID	AAA46116	standard; cDNA; 2040 BP.
XX	AAA46116;	
XX	22-AUG-2000	(first entry)
DE	Human G protein coupled receptor hGPR38 (V297K) cDNA SEQ ID NO:129	
XX	Human; G protein coupled receptor; GPCR; transmembrane receptor;	
KW	identification; agonist; screening; therapeutic; pharmaceutical;	
XX	mutant; ss.	
XX	Homo sapiens.	
OS	Synthetic.	
PN	WO200022131-A2.	
PD	20-APR-2000.	
PF	13-OCT-1999;	99WO-US24065.
XX	13-OCT-1998;	98US-0170496.
PR	12-NOV-1998;	98US-0108029.
PR	20-NOV-1998;	98US-0109213.
PR	27-NOV-1998;	98US-0110060.
PR	16-FEB-1999;	99US-0120416.
PR	26-FEB-1999;	99US-0121852.
PR	12-MAR-1999;	99US-0123944.
PR	12-MAR-1999;	99US-0123945.
PR	12-MAR-1999;	99US-0123946.

PR 12-MAR-1999; 99US-0123948.
 PR 12-MAR-1999; 99US-0123949.
 PR 12-MAR-1999; 99US-0123951.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 28-MAY-1999; 99US-0137567.
 PR 30-JUN-1999; 99US-0141448.
 PR 27-AUG-1999; 99US-0151114.
 PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156655.
 PR 29-SEP-1999; 99US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Brusma K, Chalmers DT, Chen R, Dang HT,
 PI Gore M, Law CM, Lin I, Lowitz K, White C;
 XX WPI: 2000-317986/27.
 DR P-PSDB; AAB02854.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 XX receptor, inverse or partial agonists useful as therapeutic agents
 PS Example 2; Page 166-168; 187pp; English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCRs), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA6017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 2040 BP; 350 A; 630 C; 597 G; 463 T; 0 other;

Query Match 100.0%; Score 2040; DB 21; Length 2040;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2040; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 TGCCGCCGCTCCGCGCCGCGCTGTGTCAACCGGCGCGCGCTCCGCGCTCATCGCT 480
 DB 421 TGCCGCCGCTCCGCGCCGCGCGCTGTGTCAACCGGCGCGCGCTCCGCGCTCATCGCT 480
 QY 481 GTGCTTGGGCGGTGGCGCTGCTCTTCCGCGCTCCGCTCCGCTTGGTGGGCGGTGAG 540
 DB 481 GTGCTTGGGCGGTGGCGCTGCTCTTCCGCGCTCCGCTCCGCTTGGTGGGCGGTGAG 540
 QY 541 CAGAGCCCGGCGATCTCCGATGCTCCGCGGCTCAATGAGCACCGCGCGATCGCTCTTG 600
 DB 541 CAGAGCCCGGCGATCTCCGATGCTCCGCGGCTCAATGAGCACCGCGCGATCGCTCTTG 600
 QY 601 CCTCTGCGCTGTCGCGCGCTCTCTGCTCTTCCGCGCGCGCAACCGCTCCGCGCTG 660
 DB 601 CCTCTGCGCTGTCGCGCGCTCTCTGCTCTTCCGCGCGCGCAACCGCTCCGCGCTG 660
 QY 661 GGGCGCGAGACCGCGGAGAGCGCGCGGCTGTTCACCGCGAATGCGCGCGAGCCCGCG 720
 DB 661 GGGCGCGAGACCGCGGAGAGCGCGCGGCTGTTCACCGCGAATGCGCGCGAGCCCGCG 720
 QY 721 CAGCTGGGCGGCTGCGCTGCTGATGCTGTGAGTCAACCGCGCTACTTCTTCCGCTTT 780
 DB 721 CAGCTGGGCGGCTGCGCTGCTGATGCTGTGAGTCAACCGCGCTACTTCTTCCGCTTT 780
 QY 781 CTGTGCTCAGCATCTCTTACGAGGCTCATCGGCGGAGCTGTGAGACAGCGCGCGCG 840
 DB 781 CTGTGCTCAGCATCTCTTACGAGGCTCATCGGCGGAGCTGTGAGACAGCGCGCGCG 840
 QY 841 CTGCGAGGCGCGCGCGCTCCGCGGCGGAGAGAGCGCACCGCGAGCCAAAGCGGCTG 900
 DB 841 CTGCGAGGCGCGCGCGCTCCGCGGCGGAGAGAGCGCACCGCGAGCCAAAGCGGCTG 900
 QY 901 CGTAAGTGAAGCGCGCGCTGCTTCAAAAGCGCTGCTGCACTCGCGCGCGCGGAGCC 960
 DB 901 CGTAAGTGAAGCGCGCGCTGCTTCAAAAGCGCTGCTGCACTCGCGCGCGGAGCC 960
 QY 961 GCGCAAGCGTGGTCCGCTTCCGCTGCTGCGCAAGCTGTGGGCGCGCTTCCAGCTCC 1020
 DB 961 GCGCAAGCGTGGTCCGCTTCCGCTGCTGCGCAAGCTGTGGGCGCGCTTCCAGCTCC 1020
 QY 1021 TTTTCTATTTGATTCAGCGCTCCGCGCGGCTGCTCCGCTCCGCGGAGAAACATG 1080
 DB 1021 TTTTCTATTTGATTCAGCGCTCCGCGCGGCTGCTCCGCTCCGCGGAGAAACATG 1080
 QY 1081 TCTGTGCTCCGAGAGCTCTGGGAGACCCCAAGGCGCTTTGAGGCTGAGATCCCGGAT 1140
 DB 1081 TCTGTGCTCCGAGAGCTCTGGGAGACCCCAAGGCGCTTTGAGGCTGAGATCCCGGAT 1140
 QY 1141 CGATTCAAGTAACGAGCTGCTTTCAGAGCTCTGAGACCAAGAAAGAGTGTGTAA 1200
 DB 1141 CGATTCAAGTAACGAGCTGCTTTCAGAGCTCTGAGACCAAGAAAGAGTGTGTAA 1200
 QY 1201 TTTCTTAATCCAAACCACTGTGATGAGCAAAATGAGAGGCTCTCAAGGCTCTTGAG 1260
 DB 1201 TTTCTTAATCCAAACCACTGTGATGAGCAAAATGAGAGGCTCTCAAGGCTCTTGAG 1260
 QY 1261 AGACGAGGAGATTTCAATTAAGCTAAATTTTATTAATGTAAGTATGCTGAAGC 1320
 DB 1261 AGACGAGGAGATTTCAATTAAGCTAAATTTTATTAATGTAAGTATGCTGAAGC 1320
 QY 1321 TAAAGTAAACCTTGTGCTGATCAAAAAGTAAAGATTGTGCAAGCTGTGTGAATCTT 1380
 DB 1321 TAAAGTAAACCTTGTGCTGATCAAAAAGTAAAGATTGTGCAAGCTGTGTGAATCTT 1380
 QY 1381 TTCAACAGAGAAACAGAAACTGTCTCGAAGTTGGTTTGGAGAGGAGCGCGCAAG 1440
 DB 1381 TTCAACAGAGAAACAGAAACTGTCTCGAAGTTGGTTTGGAGAGGAGCGCGCAAG 1440
 QY 1441 CGGCTTGTTCAGAGAAATGCTCTTCTGTTATGTCCAGGCTTGATTAACATATGG 1500
 DB 1441 CGGCTTGTTCAGAGAAATGCTCTTCTGTTATGTCCAGGCTTGATTAACATATGG 1500
 QY 1501 AGCTTAATGCAATTTTAAAGCAAGTATCATGACGCTGACGCTGCTCATTTTCTT 1560

Db	1501	AGCCTACTATGACGTTTAAAGCAAGATCATGACGCCGACCGCTGATCTTTTCT	1560		
Qy	1561	GGGGTGAGGATCTGCCTAGGTAGAAGTTTCTCTAATTATTTTGGCTTACTTGTTATT	1620		
Db	1561	GGGGTGAGGATCTGCCTAGGTAGAAGTTTCTCTAATTATTTTGGCTTACTTGTTATT	1620		
Qy	1621	GCAGATNGTTCCTGTCGGGGGTGGGTTATTTCGTTCCCAATGCTTTGTAAATCCC	1680		
Db	1621	GCAGATNGTTCCTGTCGGGGGTGGGTTATTTCGTTCCCAATGCTTTGTAAATCCC	1680		
Qy	1681	GGTGTCTGTCTTANGTGTGACGTGTGTGTCTTGCGCATTTATTAATTTCGGTTGCC	1740		
Db	1681	GGTGTCTGTCTTANGTGTGACGTGTGTGTCTTGCGCATTTATTAATTTCGGTTGCC	1740		
Qy	1741	TTCCACGTTGGAGATTCATTTTACATTAACACGGAAATTCGGGAGTGAATGACTTCTCT	1800		
Db	1741	TTCCACGTTGGAGATTCATTTTACATTAACACGGAAATTCGGGAGTGAATGACTTCTCT	1800		
Qy	1801	CAGTACTTTAAGATCTGCGCTCGCAACTTTTCTATCTGAGCCATCTATCAACCCCAATC	1860		
Db	1801	CAGTACTTTAAGATCTGCGCTCGCAACTTTTCTATCTGAGCCATCTATCAACCCCAATC	1860		
Qy	1861	CTCTACAACTCATTTCAAGAAGTACAGACGGCGGCTTTAACTGCTGTCTGCAGG	1920		
Db	1861	CTCTACAACTCATTTCAAGAAGTACAGACGGCGGCTTTAACTGCTGTCTGCAGG	1920		
Qy	1921	AAGTCACAGCCGAGAGGCTTCCACAAGACGAGGACACCTGCGGGGGAGTTGACGGGGAC	1980		
Db	1921	AAGTCACAGCCGAGAGGCTTCCACAAGACGAGGACACCTGCGGGGGAGTTGACGGGGAC	1980		
Qy	1981	ACTGAGAGACACGGTGGGCTTACACCCGAGACAAGCGCTAACGTTGAAGACATGGGATAA	2040		
Db	1981	ACTGAGAGAGACACGGTGGGCTTACACCCGAGACAAGCGCTAACGTTGAAGACATGGGATAA	2040		
RESULT 2					
ID	AAZ45402	AAZ45402 standard; DNA; 3066 BP.			
XX	XX				
AC	AAZ45402;				
XX	XX				
DT	27-MAR-2000	(first entry)			
XX	XX				
DE	XX	Genomic sequence of the motilin receptor gene including 5' UTR.			
XX	XX				
KM	Motilin receptor; G-protein coupled receptor; MTL-R1; GPR38;				
KM	spliced form; MTL-R1a; MTL-R1b; gastric motility disorder;				
KM	functional defect; neurological disorder; scleroderma; colonoscopy;				
KM	paraneoplastic syndrome; radiation induced dysmotility; diabetes;				
KM	infection; stress-related motility disorder; psychogenic disorder;				
KM	gastroenteritis; gastro-oesophageal reflux disease; constipation;				
KM	chronic idiopathic pseudo obstruction; acute faecal impaction;				
KM	postoperative ileus; gallstones; infantile colic; diarrhoea;				
KM	irritable bowel syndrome; non-ulcer dyspepsion; non-cardiac chest pain				
KM	endoscopy; duodenal intubation; ds.				
OS	Homo sapiens.				
XX	XX				
XX	XX				
FT	Key	Location/Qualifiers			
FT	misc_feature	1929..1931			
FT	FT	/*tag= a			
FT	FT	/*note= "imperfect donor site"			
FT	FT	1930..2728			
FT	FT	/*tag= b			
FT	FT	/*note= "intronic sequence"			
FT	FT	2080..2082			
FT	FT	/*tag= c			
FT	FT	/*note= "perfect donor site"			
FT	FT	2729..2732			
FT	FT	/*tag= d			
FT	FT	/*note= "perfect splice acceptor site"			
XX	XX				

PN	M090964436-A1.
XX	16-DEC-1999.
PF	08-JUN-1999; 99WO-US12773.
XX	
PR	12-JUN-1998; 98US-0089098.
XX	
PA	(MERI) MERCK & CO INC.
PI	Feighner SD, Patchett AA, Tan C, McKee K, MacNeil D, Howard AD;
P1	Pong S, Smith RG,
XX	WPI; 2000-105868/09.
DR	
XX	
PT	Novel receptor protein for screening compounds used in treating
PT	irritable bowel syndrome, constipation and other gastric conditions -
PS	Example 1; Fig 1; 4app; English.
XX	
CC	The present sequence represents the genomic sequence of the mottlin
CC	receptor gene, including the 5' untranslated region (5' UTR). This gene
CC	encodes a G-protein coupled receptor, and is designated MTL-R1 (also
CC	GPR38). Two spliced forms of GPR38 exist, MTL-R1A (see AAY5145) and
CC	MTL-R1B (see AAY5146). MTL-R1A is a functional seven transmembrane
CC	domain form, and MTL-R1B is a truncated five transmembrane domain. The
CC	MTL-R1 proteins are used to identify agonists and antagonists which can
CC	be used for treating gastric motility disorders, functional defects,
CC	disorders secondary to neurological disorders e.g. scleroderma,
CC	paraneoplastic syndromes radiation induced dysmotility, diabetes,
CC	infections, stress-related motility disorders, psychogenic disorders,
CC	gastroenteritis, gastro-oesophageal reflux disease, constipation,
CC	chronic idiopathic pseudo obstruction, acute faecal impaction,
CC	postoperative ileus, gallstones, infantile collic, irritable bowel
CC	syndrome, non-ulcer dyspepsion, non-cardiac chest pain and diarrhoea.
CC	They can also be used in the preparation for colonoscopy, endoscopy and
CC	duodenal intubation.
SQ	
XX	Sequence 3066 BP; 585 A; 930 C; 926 G; 625 T; 0 other;
Query Match	99.1%; Score 2021.2; DB 21; Length 3066;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2035; Conservative	0; Mismatches 3; Indels 2; Gaps 1.
OY	1 ATGGGACAGCCCCCTGTGAACGGCAGCGACGCCGCCAGGGGCGCGGAGCGCGCTGAGCCC 60
Db	1029 ATGGGCAACCCCTGTGAACGGCAGCGACGCCGCCAGGGGCGCGGAGCGCGCTGAGCCC 1088
OY	61 GGCGTGCCTTGTCGACGAGGCGCGCTGCGCCCTTCCCTGGGGGCGCTGATCCG 120
Db	1089 GGCGTGCCTTGTCGAGGAGCGCGCTGCTGCCCTTCCCTGGGGGCGCTGATCCG 1148
OY	121 GTGACCGGTGTGTGCTTGTGCTTGTCTGTCTGTGGGGGTGAGGGGCAACTGTGACCCTG 180
Db	1149 GTGACCGGTGTGTGCTTGTGCTTGTCTGTCTGTGGGGGTGAGGGGCAACTGTGACCCTG 1208
OY	181 ATGCTGATCGGAGCGCTACCGGGAACATGTGGGACAACAACAATTGTACTCTGGGACAGATG 240
Db	1209 ATGCTGATCGGAGCGCTACCGGGAACATGTGGGACAACAACAATTGTACTCTGGGACAGATG 1268
OY	241 GCCGTGTCCGACTTAATCATCTCTGTCTGGGGTCCGCGTTGCACCTGTACCCGCTCTGGCGC 300
Db	1269 GCCGTGTCCGACTTAATCATCTCTGTCTGGGGTCCGCGTTGCACCTGTACCCGCTCTGGCGC 1328
OY	301 TTGGGGGCGCCCTGGGTGTTTGGGGCGCGTGTCTGTGCGCCCTGTCCCTTAACTGTGGGCGAGGGC 360
Db	1329 TTGGGGGCGCCCTGGGTGTTTGGGGCGCGTGTCTGTGCGCCCTGTCCCTTAACTGTGGGCGAGGGC 1388
OY	361 TGCACTTACGGCCAGCGTGTCTGCACATGACCCGCGCTCACGCTGTGAGCGCTTACTTGGCCATTC 420
Db	1389 TGCACTTACGGCCAGCGTGTCTGCACATGACCCGCGCTCACGCTGTGAGCGCTTACTTGGCCATTC 1448
OY	421 TGCCGCCCGCTTCGGGCGCGCGTGTGTGTACCCGCGCCGCGGTCCGCGCGCTATCGCT 480

```

Db 1449 TGGCGGCGCGGCGCGCGCGCTTGTGTCACCGCGCGCGCGCTCATGCT 1508
Qy 481 GTGCTTGAGGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 540
Db 1509 GTGCTTGAGGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1568
Qy 541 CAGGACCCCGGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 600
Db 1569 CAGGACCCCGGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1628
Qy 601 CCTGCTGCGCTGCTGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 660
Db 1629 CCTGCTGCGCTGCTGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1688
Qy 661 GGGCGCGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 720
Db 1689 GGGCGCGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1748
Qy 721 CAGCTGGGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 780
Db 1749 CAGCTGGGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1808
Qy 781 CTGTGCTGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 840
Db 1809 CTGTGCTGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1868
Qy 841 CTGTGCGCGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 900
Db 1869 CTGTGCGCGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1928
Qy 901 CGTAACTGAGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 960
Db 1929 CGTAACTGAGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1988
Qy 961 GCGCAAAACGCTGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1020
Db 1989 GCGCAAAACGCTGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 2048
Qy 1021 TTTCCTATTTGAGTCCAGGCTCCAGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1080
Db 2049 --TTCCTATTTGAGTCCAGGCTCCAGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 2106
Qy 1081 TCCTGTGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 1140
Db 2107 TCCTGTGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCTTCTTGTCTCTGTGGGGGTGAG 2166
Qy 1141 CGATTCAGTAAACAGAGCTGCTTCCAGAGCTCTGAGACCAAGAAAGAGTGGTAA 1200
Db 2167 CGATTCAGTAAACAGAGCTGCTTCCAGAGCTCTGAGACCAAGAAAGAGTGGTAA 2226
Qy 1201 TTCTTAATCCAAACCTGTTAGATGCAAAAGTGAAGTCTCAAGTCTCTGAGA 1260
Db 2227 TTCTTAATCCAAACCTGTTAGATGCAAAAGTGAAGTCTCTGAGA 2286
Qy 1261 AGACGAGGAGATTTCTAAGCTAAATTTTATTATTAAGTAAAGATGTAAGC 1320
Db 2287 AGACGAGGAGATTTCTAAGCTAAATTTTATTATTAAGTAAAGATGTAAGC 2346
Qy 1321 TAAAGTAAACCTGCTGATCAAAAAGTAAAGTGTGACAGCTGTGTGAATCTT 1380
Db 2347 TAAAGTAAACCTGCTGATCAAAAAGTAAAGTGTGACAGCTGTGTGAATCTT 2406
Qy 1381 TTCAACGAGAAACAGAAACTTGTCTCCAGAGTGTGTGGAAGAAAGCTGCAAG 1440
Db 2407 TTCAACGAGAAACAGAAACTTGTCTCCAGAGTGTGTGGAAGAAAGCTGCAAG 2466
Qy 1441 CGGCTGTTAGAGAAATGCTCTGCTGCTTATGCTGAGCTTGAATACATATGG 1500
Db 2467 CGGCTGTTAGAGAAATGCTCTGCTGCTTATGCTGAGCTTGAATACATATGG 2526
Qy 1501 AGCTTACTATGAGTTTTAAAGCAAGTATCCATGAGCTGAGCTGTGATTTTCT 1560

```

```

Db 2527 AGCTTACTATGAGTTTTAAAGCAAGTATCCATGAGCTGAGCTGTGATTTTCT 2586
Qy 1561 GGGGTGAGATCTGCTGAGTAAAGTCTTCTAATTTATTTGCTGTTACTTGTAT 1620
Db 2587 GGGGTGAGATCTGCTGAGTAAAGTCTTCTAATTTATTTGCTGTTACTTGTAT 2646
Qy 1621 GCAGATGCTCTGCTGAGGCTGAGGCTTATTTGCTTCCATGCTTTTGTATCC 1680
Db 2647 GCAGATGCTCTGCTGAGGCTGAGGCTTATTTGCTTCCATGCTTTTGTATCC 2706
Qy 1681 GGTGCTGCTTATGTTAGAGTGTGCTGCTGCTGCTGCTTATTAATTTGCTGTTGCC 1740
Db 2707 GGTGCTGCTTATGTTAGAGTGTGCTGCTGCTGCTGCTTATTAATTTGCTGTTGCC 2766
Qy 1741 TTCAAGTGTGAGAAATCAATTTAATTAACAAGAAATTCGCGAGTATGATCTCT 1800
Db 2767 TTCAAGTGTGAGAAATCAATTTAATTAACAAGAAATTCGCGAGTATGATCTCT 2826
Qy 1801 CAGTACTTTAATGCTGCTGCTGCTGCTGCTGCTGCTTATTAATTTGCTGTTGCC 1860
Db 2827 CAGTACTTTAATGCTGCTGCTGCTGCTGCTGCTGCTTATTAATTTGCTGTTGCC 2886
Qy 1861 CTCTAACAACCTTATTAAGAAAGTAAAGAGAGGCGGCTTAACTGCTGCTGCAAG 1920
Db 2887 CTCTAACAACCTTATTAAGAAAGTAAAGAGAGGCGGCTTAACTGCTGCTGCAAG 2946
Qy 1921 AAGTCCAGGCGAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 2947 AAGTCCAGGCGAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3006
Qy 1981 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 3007 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3066

RESULT 3
AAF83684
ID AAF83684 standard; DNA; 1161 BP.
XX
XX
AC AAF83684;
XX
DT 23-JUL-2001 (first entry)
XX
DE Short form of molitin receptor, GPR-38B isoform encoding DNA.
XX
KW zsig33: signal transduction; hormone; enzyme; neural development;
KW gastric contractility; nutrient uptake; digestive; pancreatic; human;
KW insulin-like growth factor-I; growth hormone; bone; gastrointestinal;
KW glucose; osteopathic; anorectic; vulnerability; immunomodulator; GHS-R; ds;
KW G-protein coupled receptor; molitin receptor; GPR38; GPR-38B; isoform.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..1161
FT CDS /*tag= A
FT /product= "GPR-38B"
XX
EN WO200138355-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32074.
XX
PR 22-NOV-1999; 99US-0166765.
XX
PA (ZYMO ) ZYMOGENETICS INC.
XX
PI Sheppard PO, Jaspers SR, Deisher RA, Bishop PD;
XX
DR WPI; 2001-355879/37.
XX
DR P-PSDB; AAB62653.
XX

```



```

Db 448 ATGCGATCGGGCGCTACCGGAGCATGCGGACACCACTTGTACTGGGAGCATG 507
Qy 241 GCCGTGTCGACCTTACTATCTGCTGCGGAGCTGCGCTTGGACCTGTACCGGCTGAGCG 300
Db 508 GCCGTGTCGACCTTACTATCTGCTGCGGAGCTGCGCTTGGACCTGTACCGGCTGAGCG 567
Qy 301 TCGGCGCCCTGGGTGTGTTGGGCGCGCTGCTGCGCGCTGCTGCTTACGTGAGGAGG 360
Db 568 TCGGCGCCCTGGGTGTGTTGGGCGCGCTGCTGCGCGCTGCTGCTTACGTGAGGAGG 627
Qy 361 TGCACCTACGCCACGCTGCTGTGACATGACCGGCTCAGCGGTGAGCGCTTACCTGGCCATC 420
Db 628 TGCACCTACGCCACGCTGCTGTGACATGACCGGCTCAGCGGTGAGCGCTTACCTGGCCATC 687
Qy 421 TCGCGCGCGCTGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 688 TCGCGCGCGCTGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
Qy 481 GTGCTCTGGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 748 GTGCTCTGGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Qy 541 CAGGACCCCGGAGATCTCCGTAGTCCCGGAGCTCAATGACACCGCGCGGATGCGCTCTCG 600
Db 808 CAGGACCCCGGAGATCTCCGTAGTCCCGGAGCTCAATGACACCGCGCGGATGCGCTCTCG 867
Qy 601 CCTCTGCGCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 868 CCTCTGCGCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
Qy 661 GGGCGCGGAGCGCGGAGCGCGGCGCTGCTGACCGCGCATGCGCGGCGGAGCGCGCG 720
Db 928 GGGCGCGGAGCGCGGAGCGCGGCGCTGCTGACCGCGCATGCGCGGCGGAGCGCGCG 987
Qy 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 988 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047
Qy 781 CTGTGCTCAGCATCTCTCTACCGGCTCATGCGGCGGAGCTGCTGAGGAGCGCGGCGCG 840
Db 1048 CTGTGCTCAGCATCTCTCTACCGGCTCATGCGGCGGAGCTGCTGAGGAGCGCGGCGCG 1107
Qy 841 CTGCGAGCGCGCGCGCGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 1108 CTGCGAGCGCGCGCGCGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1167
Qy 901 C 901
Db 1168 C 1168

```

```

XX 10-MAY-2001.
PD 25-OCT-2000; 2000WO-US29426.
XX PF
XX PR 29-OCT-1999; 99US-0162264.
XX (MERI) MERCK & CO INC.
XX PA
XX PI Tan C, McKee K;
XX DR WPI: 2001-343479/36.
XX DR P-PSDB; AAB68477.
XX PT Novel polypeptides related to dog and rabbit motilin receptor
XX PT polypeptide, comprising unique regions from dog and motilin receptor
XX PT amino acid sequence, useful for identifying compounds for treating
XX PT diarrhoea in humans
XX PS Claim 18; Page 18-19; 42pp; English.
XX CC The present sequence encodes a rabbit motilin receptor polypeptide.
XX CC The specification describes a unique sequence present in exon 1 of
XX CC the dog motilin receptor, which is not present in human or Sphaeroides
XX CC nephelus 7557 motilin receptor sequences. The unique nucleic acid
XX CC sequence is useful for measuring the ability of a compound to affect
XX CC motilin receptor activity. Motilin receptor polynucleotides and
XX CC polypeptides are used to identify therapeutic compounds which are
XX CC useful for treating gastrointestinal diseases and disorders such as
XX CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
XX CC and diarrhoea.
XX SQ Sequence 1203 BP; 154 A; 423 C; 403 G; 223 T; 0 other;
XX
Qy 1 ATGGGCAAGCCCTTGGAGACGAGCGAGCGCCCGGAGGAGGAGCGCGCGCGCGCGCGCG 60
Db 1 ATGGGCAAGCCCTTGGAGACGAGCGAGCGCCCGGAGGAGGAGCGCGCGCGCGCGCGCG 60
Qy 61 GCGCTGCGCGCTTGGAGAGCGCGCGCTGCGCGCTTCCCTGAGGAGCGCGCGCGCGCGCG 120
Db 61 GCGCTGCGCGCTTGGAGAGCGCGCGCTGCGCGCTTCCCTGAGGAGCGCGCGCGCGCGCG 120
Qy 121 GTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 GTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy 181 ATGCTGATCGGCGCTGACCGGAGCATGCGGACCAACCACTTGTACTGGGAGCATG 240
Db 181 CTGCTGATCGGCGCTGACCGGAGCATGCGGACCAACCACTTGTACTGGGAGCATG 240
Qy 241 GCCGTGTCGACCTTACTATCTGCTGCGGAGCTGCGCTTGGACCTGTACCGGCTGAGCG 300
Db 241 GCCGTGTCGACCTTACTATCTGCTGCGGAGCTGCGCTTGGACCTGTACCGGCTGAGCG 300
Qy 301 TCGGCGCCCTGGGTGTGTTGGGCGCGCTGCTGCGCGCTGCTGCTTACGTGAGGAGG 360
Db 301 TCGGCGCCCTGGGTGTGTTGGGCGCGCTGCTGCGCGCTGCTGCTTACGTGAGGAGG 360
Qy 361 TGCACCTACGCCACGCTGCTGTGACATGACCGGCTCAGCGGTGAGCGCTTACCTGGCCATC 420
Db 361 TGCACCTACGCCACGCTGCTGTGACATGACCGGCTCAGCGGTGAGCGCTTACCTGGCCATC 420
Qy 421 TCGCGCGCGCTGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 TCGCGCGCGCTGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 GTGCTCTGGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 GTGCTCTGGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

```

QY 541 CAGACCCCGCATCTCCGTAAGTCCCGGACCTTCATAGCACCGCGCGATCGCTCTCTCG 600
 Db 541 CAGACCCCGCGCTCTTCGCGGACCCCGGACCGCAACGTAATGTGCGCTGACCCCTCG 600
 QY 601 CCTTCGCGCTGTGCGCGGCTCTTCGCGCTCTCGCGGCGCGACCGCGCTCTCGCGCTCG 660
 Db 601 TCGCCCGCC-----CCGCGCTCTCGCGCTCG 627
 QY 661 GGGCCCGACCGCGAGCGCGCGCGCTGTTACGCGCGAATGCGCGCGCGCGCGCG 720
 Db 628 GGGCCCG---GAGCGGAGCGCGCGCGCTGTTACGCGCGAAGTCCGCGCGCGCGCGCG 664
 QY 721 CAGCTGGGCGCGCTGCGGTCTCATGCTGTGGGTACCAACCGCTCTCTCTCGCGCTTT 780
 Db 685 CAGCTGGGCGCTGCGGTCTCATGCTGTGGGTACCAACCGCTCTCTCTCGCGCTTT 744
 QY 781 CTGCGCTCAGCATCTCTACGCGGCTCATGCGGCGGAGACTGTGGAGCAGCGCGCGCG 840
 Db 745 CTGCGCTCAGCATCTCTACGCGGCTCATGCGGCGGAGACTGTGGAGGCTCGGCGCGCG 804
 QY 841 CTGCGAGCGCGCGCGCGCTCGGCGGCGGAGAGAGCGCACCGGACCAACCGCGCTCG 900
 Db 805 CTGCGAGCGCGCGCGCGCGCGCTCGGCGGAGAGAGCGCACCGGACCAACCGCGCTCG 864
 QY 901 C 901
 Db 865 C 865

RESULT 10

AAf85447 ID AAF85447 standard; DNA; 813 BP.

AC AAF85447;

DT 23-JUL-2001 (first entry)

XX Nucleotide sequence of a dog motilin receptor exon 1.

XX Motilin receptor; gastrointestinal disease; gastric motility disorder;

KM gastroparesis; irritable bowel syndrome; diarrhoea, ss.

XX Canis sp.

XX Key Location/Qualifiers

FT CDS 1..813

FT /+tag= a

FT /note= "motilin receptor exon 1"

XX MO200132710-A1.

XX 10-MAY-2001.

XX 25-OCT-2000; 2000MO-US29426.

XX 29-OCT-1999; 99US-0162264.

XX (MERI) MERCK & CO INC.

XX Tan C, McKee K;

XX WPI, 2001-343479/36.

XX P-PSDB; AAB68476.

XX Novel polypeptides related to dog and rabbit motilin receptor

XX polypeptide, comprising unique regions from dog and motilin receptor

XX amino acid sequence, useful for identifying compounds for treating

XX diarrhoea in humans

XX Claim 14; Page 17-18; 42pp; English.

XX The present sequence represents exon 1 of a dog motilin receptor gene.

CC The specification describes an unique sequence present in exon 1 of
 CC the motilin receptor, which is not present in human or Spharoides
 CC nephelus 75E7 motilin receptor sequences. The unique nucleic acid
 CC sequence is useful for measuring the ability of a compound to affect
 CC motilin receptor activity. Motilin receptor polynucleotides and
 CC polypeptides are used to identify therapeutic compounds which are
 CC useful for treating gastrointestinal diseases and disorders such as
 CC gastric motility disorders, gastroparesis, irritable bowel syndrome,
 CC and diarrhoea.

SQ Sequence 813 BP; 65 A; 316 C; 305 G; 127 T; 0 other;

Query Match 26.5%; Score 541.2; DB 22; Length 813;

Best Local Similarity 80.1%; Pred. No. 2.1e-101;

Matches 702; Conservative 0; Mismatches 108; Indels 66; Gaps 3;

QY 25 GAGGAGCCCGAGGAGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 84
 Db 4 GAGGAGCCCGAGGAGGAGCGCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCG 63
 QY 85 CGCTGCTCGCGCTTTCCCTGGGAGGCGCTGTGCGCGGAGACCGGTCGTGCTGCGCTG 144
 Db 64 CTGTGCTCGCGCTTTCCCTGGGAGGCGCTGTGCGCGGAGACCGGTCGTGCTGCGCTG 123
 QY 145 TTGCTGCTCGGAGTGAAGCGGCAACGTGTGACCTGTGATCGAGCGCTACCGGAGC 204
 Db 124 TTGCGCGTGGCGGTGAAGCGGCAACGTGTGACCTGTGATCGAGCGCTACCGGAGC 183
 QY 205 ATGCGAGCACCAACCACTTGTACTGTGGAGACATGACCGCTGTGCACTATCTGCTG 264
 Db 184 ATGCGAGCACCAACCAACCTGTACTGTGGAGACATGACCGCTGTGCACTATCTGCTG 243
 QY 265 CTGAGGCGTGGTTCACCTGTACCGCGCTCGGCGCGCTCGGCGCGCGCGCGCGCG 324
 Db 244 CTGAGGCGTGGTTCACCTGTACCGCGCTCGGCGCGCTCGGCGCGCGCGCGCGCG 303
 QY 325 CTGCTGTCCGCGCTGTCTCTACGTGTGGAGAGGAGCTGCACTAGCGACGCTGTGAC 384
 Db 304 CTGCTGTCCGCGCTGTCTCTACGTGTGGAGAGGAGCTGCACTAGCGACGCTGTGAC 363
 QY 385 ATACCGCGCTACGCTGTGAGCGCTTACCTGTGCGACATCTGCGCGCGCGCGCGCTC 444
 Db 364 GTACCGCGCTGAGCGCTGAGCGCTTACCTGTGCGCGCGCGCGCGCGCGCGCG 423
 QY 445 TTGCTACCGCGCGCGCGCGCGCGCGCTCATGCTGTGCTGTGGAGCGTGTGCTGTC 504
 Db 424 CTGCTGTCCGCGCGCGCGCGCGCGCGCTCATGCGCGCGCTGTGGAGCGCTGTGTC 483
 QY 505 TCTGCGGATCCCTTCTTGTCTGTGTGAGCGGTGAGCAGGACCGCGCATCTCCGTAGTC 564
 Db 484 TCGGCGGCGCGCTTCTTCTTCTGTGTGAGCGGTGAGCAGGACCGCGCATCTCCGTAGTC 534
 QY 565 CGGAGCGCTCAATGAGCACCGGCGGAGATGCGCTCTCGCTCTGCGCTGTGCGCGCTTC 624
 Db 535 CCGGCGCTCAACGAGCAGCGCGG----- 557
 QY 625 TGAGCTTCGCGGAGCGCACCGCGCTCCCGCGCGTGGAGGCGCGAGACCGGAGCGCGG 684
 Db 558 -----GCTGAGCGGAGCGCGCTCCCGCGCGCGCGCGCGG-----GCG 597
 QY 685 GCGCTGTTCAGCGCGAATGCGCGCGAGCGCGCGAGCTGTGAGCGCGCTGTGATG 744
 Db 598 GCGCTTTCAGCGGAGAGTGGCGGCGAGCGCGCTGTGAGCTGTGAGCGCGCTGTGATG 657
 QY 745 CTGTGAGTACCAACCGCTTCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 804
 Db 658 CTGTGAGTACCAACCGCTTCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 717
 QY 805 CTCATGAGCGGAGCTGTGAGCAGCGCGCGCGCTGTGAGAGCGCGCGCGCTGTGAGG 864
 Db 718 CGCATTCGCGCGAGCTGTGAGAGCGCGCGCGCGCTGTGAGAGCGCGCGCGCTGTGAGG 777
 QY 865 CGGAGAGAGCGCACCGGAGAGCAACCAACCGCTGCTG 900

PR 05-SEP-2000; 2000DE-1044543.
 PA (EPiG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ1310-ABQ4121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 XX Sequence 1179 BP; 399 A; 443 C; 159 G; 178 T; 0 other;
 SQ
 Query Match 24.6%; Score 501.2; DB 24; Length 1179;
 Best Local Similarity 79.9%; Pred. No. 3.4e-93;
 Matches 590; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 481 GTGCTTGGGCGGCTGCTCTGTCGCGATCCCTTCTTGTTCGTGGAGCGTCGAG 540
 DB 922 ATACTTAACCCGTAAGCTACTCTCTACCGATCCCTTCTTATCTTAATAAGCTCGAA 981
 QY 541 CAGAACCCCGGATCTCCGTAGTCCCGGCTCAATGACACCGCGGATGCTCTCTCG 600
 DB 982 CAAACCCCGACATCTCCGAATCCGAACTCAATACACCGCGGATGCTCTCTCG 1041
 QY 601 CCTCTGCTCTGCTGCGCGCTCTGCTGCTGCGGCGGCGACCGCGTCCCGCGCTCG 660
 DB 1042 CTTCTGCTCTGCTGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
 QY 661 GGGCGGAGACCGCGGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 DB 1102 AAACCCGAAACCGCGGAAACCGCGACGCTATTCAACCGGAAATCCGACCGAACCCCGG 1161
 QY 721 CAGCTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
 DB 1162 CAACTAAACCGGCTACGT 1179

RESULT 13
 ABQ47148
 ID ABQ47148 standard; DNA; 1179 BP.
 XX
 AC ABQ47148;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 33739.
 XX
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 PN W0200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 PA (EPiG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide

QY 241 GCCGCTCCGACCTACTATCTCTCTGCGGCTGCGCTGACCTGATACCGCCTCTGCGGC 300
 CC 498 GTCGCTGCTGATTTATTTATTTGTCGGGTCGTGCTGATTTGATCGTTTGGCGGT 439
 QY 301 TCGGCGCCCTGGGTCGTCGGGCGCGGCTCTGCGGCTGCTGCTGATACCGGCGAGGC 360
 Db 438 TCGCGGTTTGGGTCGTCGGGTCGTCGTTTGTGCTTTGTTTATTCGTCGGGCGAGGCT 379
 QY 361 TGCACCTACGACCGCTGCTGACATGACCGGCTCAGGCTCAGGCTACTGAGCCATC 420
 Db 378 TGTATTACGTTACGTTGTTATATGATCGGCTTACGCTGACGCTATTTGCTATT 319
 QY 421 TCGCGCCGCTCTCGCGCCGCGCTTGTGTCACCGGCGCGGCTCGCGGCTCATGCT 480
 Db 318 TGTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
 QY 481 GTGCTCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 258 GTGCTTGGGTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199
 QY 541 CAGGACCCCGGCTATCCGTCAGTCCGCGGCTCATATGACGCGCGGATCCGCTCTCG 600
 Db 198 TAGGATTCGCGATTTGCTGATTTGCGGTTTAAATGATCGCGCGGATCGTTTTCG 139
 QY 601 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 Db 138 TTTTTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79
 QY 661 GGGGCGGAGACCGGAGGCGCGGCTGCTCAGCGCGGATCCGCGCGGAGCCCGCG 720
 Db 78 GGGTTCGAGATCGCGGAGGTCGCGGCTGTTTAAATGCGGATGCGGATGCTTTCG 19
 QY 721 CAGCTGCGCGGCTGCTGCT 738
 Db 18 TAGTTGGGCGGCTGCT 1

RESULT 15
 AAV4930
 ID AAV4930 standard; cDNA: 283 BP.
 XX
 AC AAV44930;
 XX
 DT 24-NOV-1998 (first entry)
 XX
 DE Galanin receptor GALR2 DNA probe.
 XX
 KM Galanin receptor; GALR2; rat; ligand; obesity; anorexia; pain;
 XX
 KM cognitive disorder; therapy; probe; ss.
 XX
 OS Rattus sp.
 XX
 PN W09829441-A1.
 XX
 PD 09-JUL-1998.
 XX
 PF 18-DEC-1997; 97WO-US23892.
 XX
 PR 27-DEC-1996; 96US-0033851.
 XX
 PA (MERT) MERCK & CO INC.
 XX
 PA (UYTE-) UNIV TEXAS HEALTH SCI CENT SAN ANTONIO.
 XX
 PA (UTOR) UNIV TORONTO.
 XX
 PI Kolakowski LF, O'Dowd B, Sullivan K;
 DR WPI; 1998-388039/33.
 XX
 PT New human galanin receptor, GALR2, - useful to identify agonists and
 PT antagonists to treat conditions involving galanin, e.g. for
 PT treatment of obesity or cognitive disorders
 XX

PS Example 1; Fig 6; 57bp; English.

CC This PCR fragment was used as a probe to screen a rat hypothalamus
 CC cDNA library. 2 independent clones, named 27A (see AAV4929) and
 CC 16.6, were obtained. Clone 27A codes for a novel full-length rat
 CC galanin receptor, designated GALR2 (see AAV61385). The invention
 CC provides methods for identifying ligands particular to human GALR2
 CC (see AAV61386). Such ligands may be useful therapeutically e.g. to
 CC treat obesity or cognitive disorders involving excess galanin or to
 CC treat pain or anorexia involving insufficient galanin.
 XX

SQ Sequence 283 BP; 27 A; 116 C; 84 G; 56 T; 0 other;

Query Match 13.9%; Score 283; DB 19; Length 283;
 Best Local Similarity 100.0%; Pred. No. 1.1e-48;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 TCGGACACACCACTTGTACTGAGGACAGATGCCGTGTCCGACTCATCTGTC 265
 Db 1 TCGGACACACCACTTGTACTGAGGACAGATGCCGTGTCCGACTCATCTGTC 60
 QY 266 TCGGACACACCACTTGTACTGAGGACAGATGCCGTGTCCGACTCATCTGTC 325
 Db 61 TCGGACACACCACTTGTACTGAGGACAGATGCCGTGTCCGACTCATCTGTC 120
 QY 326 TGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
 Db 121 TGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 386 TGCACCGGCTCAGGCTGAGGCGCTACTGCGCATCTGCGCGCGCTCCGCGCGTCT 445
 Db 181 TGCACCGGCTCAGGCTGAGGCGCTACTGCGCATCTGCGCGCGCTCCGCGCGTCT 240
 QY 446 TGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
 Db 241 TGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 283

Search completed: March 16, 2003, 01:21:47
 Job time : 489 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 10:25:58 ; Search time 5959 Seconds

(without alignments)
9963.035 Million cell updates/sec

Title: US-09-876-252-129

Perfect score: 2040
Sequence: 1 atgggagagcccccgaacgg.....acgtgaagacgatggatata 2040

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pac: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rod: *
36: em_hcg_man: *
37: em_hcg_vtc: *
38: em_sy: *
39: em_hcg_hum: *
40: em_hcg_mus: *
41: em_hcg_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2035.2	99.8	2040	9	AF034632	AF034632 Homo sapi
2	2033.6	99.7	163284	9	AL137000	AL137000 Human DNA
3	2033.6	99.7	341560	2	AL596304	AL596304 Homo sapi
4	1047.2	51.3	1161	6	AX154591	AX154591 Sequence
5	896.2	43.9	1239	6	AX154589	AX154589 Sequence
6	592.2	29.0	692	9	HS339407	AJ339407 Homo sapi
7	572.2	28.0	692	9	HS342408	AJ342408 Homo sapi
8	557.4	27.3	615	9	HS339459	AJ339459 Homo sapi
9	424.8	20.8	701	9	HS3326768	AJ3326768 Homo sapi
10	307	15.0	172168	2	AC109158	AC109158 Mus muscu
11	283	13.9	283	6	AR168467	AR168467 Sequence
12	283	13.9	283	6	AR182284	AR182284 Sequence
13	266.8	13.1	573	6	HS341295	AJ341295 Homo sapi
14	234.8	11.5	1063	6	AR156351	AR156351 Sequence
15	234.8	11.5	1101	4	SSU60178	U60178 Sus scrofa
16	233.2	11.4	870	4	SSU60180	U60180 Sus scrofa
17	230	11.3	1029	6	AR156352	AR156352 Sequence
18	230	11.3	142736	2	AC117349	AC117349 Rattus no
19	230	11.3	169279	2	AC095173	AC095173 Rattus no
20	230	11.3	181578	2	AC130771	AC130771 Rattus no
21	230	11.3	195125	2	AC108400	AC108400 Mus muscu
22	230	11.3	227035	2	AC116585	AC116585 Mus muscu
23	228.8	11.2	870	9	HSU60181	U60181 Human growt
24	228.8	11.2	1122	6	AR156354	AR156354 Sequence
25	228.8	11.2	6787	9	AF369786	AF369786 Homo sapi
26	228.8	11.2	145593	9	AC069523	AC069523 Homo sapi
27	228.4	11.2	1092	6	AR156357	AR156357 Sequence
28	228.4	11.2	1350	10	AB001982	AB001982 Rattus no
29	228.4	11.2	3129	6	AR156356	AR156356 Sequence
30	228.4	11.2	3129	10	RNU94321	U94321 Rattus norv
31	227.8	11.2	1088	6	AR156353	AR156353 Sequence
32	227.8	11.2	1101	6	AX154584	AX154584 Sequence
33	227.8	11.2	1101	9	HSU60179	U60179 Human growt
34	225.2	11.0	1676	5	AF082210	AF082210 Spheroeid
35	221	10.8	560	10	AY056474	AY056474 Mus muscu
36	216	10.6	1233	4	AY093950	AY093950 Ovis arie
37	216	10.6	1591	4	AY093949	AY093949 Ovis arie
38	216	10.6	1816	4	AY093948	AY093948 Ovis arie
39	214	10.5	1455	5	AF082209	AF082209 Spheroeid
40	194.4	9.5	557	10	AF332997	AF332997 Mus muscu
41	184.6	9.0	1098	6	AX154586	AX154586 Sequence
42	176.8	8.7	528	11	G68328	G68328 GHR X2.3 H
43	161.4	7.9	130311	2	AC095360	AC095360 Rattus no
44	160.4	7.9	130311	2	AC095360	AC095360 Rattus no
45	153	7.5	836	6	AR156355	AR156355 Sequence

ALIGNMENTS

RESULT 1	AF034632	2040 bp	DNA	linear	PRI 13-JUL-1998
LOCUS	AF034632				
DEFINITION	Homo sapiens orphan G protein-coupled receptor (GPR38) gene,				
ACCESSION	AF034632				
VERSION	AF034632.1	GI:2654158			
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE					
AUTHORS	1 (bases 1 to 2040)				
	McKee,K.K., Tan,C.P., Palya,O.C., Liu,J., Feighner,S.D.,				
	Hreniuk,D.L., Smith,R.G., Howard,A.D. and Van der Ploeg,L.H.				

TITLE Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors

JOURNAL Genomics 46 (3), 426-434 (1997)

MEDLINE 98110578

PUBMED 9441746

AUTHORS 2 (bases 1 to 2040)

McKee, K.K., Tan, C.P., Palyha, O.C., Liu, J., Feighner, S.D., Hreniuk, D.L., Smith, R.G., Van Der Ploeg, L.H.T. and Howard, A.D.

TITLE Direct Submission

JOURNAL Submitted (17-NOV-1997) Biochemistry and Physiology, Merck and Co., Inc., PO Box 2000, Rahway, NJ 07065, USA

REFERENCE Location/Qualifiers

FEATURES

source 1..2040

organism="Homo sapiens"

db_xref="taxon:9606"

chromosome="13"

map="13q14-21"

<1..>2040

gene /gene="GPR38"

mRNA join(<1..901,1703..>2040)

cds /gene="GPR38"

join(1..901,1703..2040)

/note="Orphan G protein-coupled receptor related to growth hormone secretagogue and neurotensin receptors"

/codon_start=1

/product="GPR38"

/protein_id="AAC26081.1"

/db_xref="GI:2654159"

/translation="MGSPWNGSDPDEGAREPPALPCDERCSPEPLGALVPTAV
CICLPVGVSGNVVTVMLIGRYDMRTTTLVYGSMAVSPLLILGLPDLVLRSR
PVMVGRGLRLSLVYVGGCTATLHMTALSVRYALICPLARVLYTRRVRLALIA
VLMVAALLSAGPLFLVGVQDPSISVPLNGSTARSPLASPLMSPPSP
PSGPETAALALFSLRECRSPADQGLRVLMVMTATVTFPLCLSLTYELIGRELS
SRPLPGAPASGRGRHQRVRLVLAFLICWLPFHVGRLLIYINTEDSRMVYFSQ
YENVALQFLYLSASINPLIYNLISKVYRAAPFLILARKSRPRGFRSHDIAVEVAG
DTGGTVGTETSAVNTMG"

BASE COUNT 347 a 631 c 598 g 464 t

ORIGIN

Query Match 99.8%; Score 2035.2; DB 9; Length 2040;

Best Local Similarity 99.3%; Pred. No. 1e-304;

Matches 2037; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 421 TGCCGCCGCTCCGCGCCGCGCTTGTGTCACCCGCGCCGCGCTCCGCGCTCATCGCT 480

DB 421 TGCCGCCGCTCCGCGCCGCGCTTGTGTCACCCGCGCCGCGCTCCGCGCTCATCGCT 480

QY 481 GTGCTTGGGCGCGTGGCGTCTCTTCCGCGTCCCTCTTGTCTGGTGGGCGTCCAG 540

DB 481 GTGCTTGGGCGCGTGGCGTCTCTTCCGCGTCCCTCTTGTCTGGTGGGCGTCCAG 540

QY 541 CAGAACCCCGCATCTCCGATGCCGCGGCTCAATGAGCACCGCGAGATCCCTCTCG 600

DB 541 CAGAACCCCGCATCTCCGATGCCGCGGCTCAATGAGCACCGCGAGATCCCTCTCG 600

QY 601 CCTCTGCGCTGCGCGCTCTCTGCTCTTCCGCGCGCGCACCGCGTCCCGCGCTG 660

DB 601 CCTCTGCGCTGCGCGCTCTCTGCTCTTCCGCGCGCGCACCGCGTCCCGCGCTG 660

QY 661 GGGCCCGAGACCGCGAGAGCGCGCGCGCTGTACCGCGAATGCGCGCGAGCCCGCG 720

DB 661 GGGCCCGAGACCGCGAGAGCGCGCGCGCTGTACCGCGAATGCGCGCGAGCCCGCG 720

QY 721 CAGCTGGCGCGCTGCTGTGATGCTGTGAGTCAACACCGCTACTTCTTCCGCTTT 780

DB 721 CAGCTGGCGCGCTGCTGTGATGCTGTGAGTCAACACCGCTACTTCTTCCGCTTT 780

QY 781 CTGTGCTCTACGATCTCTTACGAGGCTCATCGCGCGGAGCTGTGAGCACCGCGCG 840

DB 781 CTGTGCTCTACGATCTCTTACGAGGCTCATCGCGCGGAGCTGTGAGCACCGCGCG 840

QY 841 CTGCGAGCGCGCGCGCGCTCGCGCGGAGAGAGGCAACCGGACAGCCAAACGCTCT 900

DB 841 CTGCGAGCGCGCGCGCGCTCGCGCGGAGAGAGGCAACCGGACAGCCAAACGCTCT 900

QY 901 CGTAAGTGAAGCGCGCGCTGTGTCCTCAAGACGCTGCTGCGCGCGCGCGGAGAC 960

DB 901 CGTAAGTGAAGCGCGCGCTGTGTCCTCAAGACGCTGCTGCGCGCGCGCGGAGAC 960

QY 961 GCGCAACGCTGGTGGCTTCCCTCTGCTGCTGCGCGAGCTTGGGCGCGCTTCAAGTCC 1020

DB 961 GCGCAACGCTGGTGGCTTCCCTCTGCTGCTGCGCGAGCTTGGGCGCGCTTCAAGTCC 1020

QY 1021 TTTCTATTTTGGATTCAGGCTCCACCGCGCGGATCTTCCATCCCGGAGAAACATG 1080

DB 1021 TTTCTATTTTGGATTCAGGCTCCACCGCGCGGATCTTCCATCCCGGAGAAACATG 1080

QY 1081 TCCTGTCCCGCAGAGCTCTGGGAGACCCAGAGCGCTTTGAGGATGAGATCCCGGATC 1140

DB 1081 TCCTGTCCCGCAGAGCTCTGGGAGACCCAGAGCGCTTTGAGGATGAGATCCCGGATC 1140

QY 1141 CGATTCAATACCAAGTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGATGGTAA 1200

DB 1141 CGATTCAATACCAAGTCTTTTTCAGAGCTCTGAGACCAAGAAAGAGATGGTAA 1200

QY 1201 TTCTTAATCCACCACTTGTGATGAGCCAAATGAGAGCTCTTCAAGTCTTGAAG 1260

DB 1201 TTCTTAATCCACCACTTGTGATGAGCCAAATGAGAGCTCTTCAAGTCTTGAAG 1260

QY 1261 AGACGAGGAGATTCATTAAGTAAATTTTATTAATGATTAATGATGATGATGAGG 1320

DB 1261 AGACGAGGAGATTCATTAAGTAAATTTTATTAATGATTAATGATGATGATGAGG 1320

QY 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTTGTGAGACCTGTTGATAATTTCT 1380

DB 1321 TAAAGTAAACCTTGTCTGATCAAAAAGTAAAGATTTGTGAGACCTGTTGATAATTTCT 1380

QY 1381 TTCAACAGAGAACAGAAAACCTTGTGAGAGGTTTGTGAGAGGAGGAGGAGGAGG 1440

DB 1381 TTCAACAGAGAACAGAAAACCTTGTGAGAGGTTTGTGAGAGGAGGAGGAGGAGG 1440

QY 1441 CGGCTTTGTCAGAGAAATTTCTTCTGTTATGATGATGATGATGATGATGATGATG 1500

DB 1441 CGGCTTTGTCAGAGAAATTTCTTCTGTTATGATGATGATGATGATGATGATGATG 1500

QY 1501 AGCTTAATGAGATTTTAAAGCAAGTATCATGAGCTGAGCTGAGCTGATCTTTTCT 1560

```

Db 1501 ACCCTACTATGCACTTTTAAAGCAAGATCCAGACGCCGAGCCGCGATCTTTTCT 1560
QY 1561 GGGGTGAGAGATCGCTAGAGTAAAGTTCTCTAAATTTTGTGCTACTGTATTT 1620
Db 1561 GGGGTGAGAGATCGCTAGAGTAAAGTTCTCTAAATTTTGTGCTACTGTATTT 1620
QY 1621 GCAGATGTTCTCTGTCGGGGGTGGGGGGTTTATTGCTGCCAATGCTTTGTTATCCC 1680
Db 1621 GCAGATGTTCTCTGTCGGGGGTGGGGGGTTTATTGCTGCCAATGCTTTGTTATCCC 1680
QY 1681 GATGCTGCTCTTATGTTGTCAGTGTGTGTTCTGSCATTTTAAATTTGCTGTTGCC 1740
Db 1681 GATGCTGCTCTTATGTTGTCAGTGTGTGTTCTGSCATTTTAAATTTGCTGTTGCC 1740
QY 1741 TTCCAGTTGGAGATCATTTTACATMAACAGGAAGATTCGGGATGATGATCTCTCT 1800
Db 1741 TTCCAGTTGGAGATCATTTTACATMAACAGGAAGATTCGGGATGATGATCTCTCT 1800
QY 1801 CAGTACTTTAATAGTGTGCTCTGCACTTTTCTATCTGAGCCATCTTAAACCCATC 1860
Db 1801 CAGTACTTTAATAGTGTGCTCTGCACTTTTCTATCTGAGCCATCTTAAACCCATC 1860
QY 1861 CTCTCAACCTCATTTTCAAGAAGTACAGAGCGCGGCTTTAACTGCTGCTCGCAAG 1920
Db 1861 CTCTCAACCTCATTTTCAAGAAGTACAGAGCGCGGCTTTAACTGCTGCTCGCAAG 1920
QY 1921 AAGTCCAGGCGGAGAGGCTTCCACAGAGCAGAGCACTGCGGGGAGTTGACAGGGGAC 1980
Db 1921 AAGTCCAGGCGGAGAGGCTTCCACAGAGCAGAGCACTGCGGGGAGTTGACAGGGGAC 1980
QY 1981 ACTGAGAGAGACACGGTGGGCTTACACCCGACAGAGCGCTTAACTGAAAGCATGGATTA 2040
Db 1981 ACTGAGAGAGACACGGTGGGCTTACACCCGACAGAGCGCTTAACTGAAAGCATGGATTA 2040

```

RESULT 2

AL137000

LOCUS

DEFINITION

AL137000 163284 bp DNA linear PRI 18-DEC-2000
 Human DNA sequence from clone RP11-203116 on chromosome 13 contains
 the gene for KIAA0970 protein, COX7CP1 (cytochrome c oxidase
 subunit VIc pseudogene 1), a novel pseudogene, the GPR38 (G
 protein-coupled receptor 38) gene, ESTs, STSs, GSSs and a Cpg
 island, complete sequence.

ACCESSION

AL137000

VERSION

AL137000.6

KEYWORDS

HG; COX7CP1; Cpg island; cytochrome c oxidase; G protein-coupled
 receptor; GPR38; KIAA0970.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS

Wall, M.

TITLE

Direct Submission

JOURNAL

Submitted (18-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CE10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk

COMMENT

On Aug 29, 2000 this sequence version replaced gi:9926419.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accessions
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMBEP; Information
 on the WORMBEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 13, constructed by the Sanger Centre Chromosome 13
 Mapping Group. Further information can be found at

FEATURES

source

<http://www.sanger.ac.uk/HGP/Chr13>
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-203116 it may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP11-452110 is at 1 in this sequence.
 The true right end of clone RP11-103118 is at 163284 in this
 sequence. This sequence has been finished according to sequence map
 criteria as follows. An attempt is made to resolve all sequencing
 problems, such as compressions and repeats, but not necessarily
 within known annotated repeat sequence elements. Where the
 sequence is ambiguous there is an annotation using the 'unseq'
 feature key. RP11-203116 is from the library RP11-11.1 constructed
 by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6.

location/Qualifiers

1..163284	/organism="Homo sapiens"
/db_xref="taxon:9606"	
/chromosome="13"	
/clone_11b="RP11-203116"	
/clone_11b="RP11-11.1"	
3..121	
/note="MER20 repeat: matches 4..126 of consensus"	
152..261	
/note="MIR repeat: matches 20..137 of consensus"	
262..554	
/note="AluX repeat: matches 1..293 of consensus"	
555..698	
/note="MIR repeat: matches 137..255 of consensus"	
1012..1291	
/note="AluJ repeat: matches 7..291 of consensus"	
1320..1425	
/note="MER94 repeat: matches 4..114 of consensus"	
1591..1804	
/note="L1P8 repeat: matches 5945..6158 of consensus"	
3240..3621	
/note="L2 repeat: matches 2310..2691 of consensus"	
3824..4106	
/note="AluJ repeat: matches 1..301 of consensus"	
4975..5042	
/note="L2 repeat: matches 2639..2705 of consensus"	
6113..6408	
/note="AluX repeat: matches 1..295 of consensus"	
7646..7732	
/note="AluSc repeat: matches 1..85 of consensus"	
9887..10191	
/note="AluX repeat: matches 1..306 of consensus"	
12201..12247	
/note="L2 repeat: matches 2704..2750 of consensus"	
12472..12738	
/note="L2 repeat: matches 1304..1530 of consensus"	
12739..12874	
/note="L2 repeat: matches 2..133 of consensus"	
12875..12887	
/note="L2 repeat: matches 1530..1542 of consensus"	
12922..13227	
/note="L2 repeat: matches 5871..6183 of consensus"	
13303..13348	
/note="MIR repeat: matches 1..46 of consensus"	
13533..13623	
/note="MIR repeat: matches 223..315 of consensus"	
13657..13865	
/note="MIR repeat: matches 307..533 of consensus"	
13890..14150	
/note="MIR repeat: matches 195..456 of consensus"	
14151..14450	
/note="AluX repeat: matches 1..307 of consensus"	
14501..14729	
/note="L1M4C repeat: matches 588..813 of consensus"	
14730..15006	
/note="AluJ repeat: matches 1..285 of consensus"	
15007..15361	
repeat_region	

Db 321704 ACTGAGAGAGACCGTGGGCTTACCGAGACAGCGCTTACGTAAGAGATGAGATAA 321763

RESULT 4
AX154591 1161 bp DNA linear PAT 23-JUN-2001

LOCUS
DEFINITION Sequence 11 from Patent WO0138355.
ACCESSION AX154591
VERSION AX154591.1 GI:14536177
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1161)
Sheppard, P.O., Jaspers, S.R., Deisher, T.A. and Bishop, P.D.
Method of forming a peptide-receptor complex with zsig33 and
therapeutic use thereof
Patent: WO 0138355-A 11 31-MAY-2001;
ZymoGenetics, Inc. (US)
FEATURES
Location/Qualifiers
1..1161
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1161
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC42612.1"
/db_xref="GI:14536178"
/translation="MGSPMNGSDGPEGAREPPWALPPEDRRCSPPFLGALVPTAV
CLCLFVVGSGNVTVMLIGRYRDMRTTNILGSMAYSDLLILGLPFDLYRWRSR
PWFPLICRSLIVGSGCTATLLHMTALSVERYALICPLRARVLVTRRRVALIA
VMAVALISAGPFLFIVGVEDPDISIVPGLNGLARIASPLASPPMLSRAPPS
PSGSETLEAALFSECRPSPAQGLALRVMVMTAYAFELPCLISILYGLRGLWS
SRRLRGPASGRERGRQTVLRKMSRSGKXACIOSAPPGTACGLPILLAQLM
APLPAPPISPISAPRRGGSGIYNLVALPRMNHKHFADVLISLV"

BASE COUNT 126 a 436 c 372 g 227 t

ORIGIN

Query Match 51.3%; Score 1047.2; DB 6; Length 1161;
Best local Similarity 99.7%; Pred. No. 3.6e-153;
Matches 1049; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGCGAGCCCTTGGAAACGAGCGAGCGCCCGAGGAGGCGGAGCGCGCTGGCC 60
Db 1 ATGGGCGAGCCCTTGGAAACGAGCGAGCGCCCGAGGAGGCGGAGCGCGCTGGCC 60

QY 61 GCGGTGCGGCTTGGAGAGCGGCGGCTGCGCCCTTCCCTGGGGCGCTGGTGC 120
Db 61 GCGGTGCGGCTTGGAGAGCGGCGGCTGCGCCCTTCCCTGGGGCGCTGGTGC 120

QY 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180
Db 121 GTGACCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180

QY 181 ATGCTGATCGGGGCTTACCGGAGCATGCGGACCAACCACTTGTACTGGGACGATG 240
Db 181 ATGCTGATCGGGGCTTACCGGAGCATGCGGACCAACCACTTGTACTGGGACGATG 240

QY 241 GCGGTGCGGAGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GCGGTGCGGAGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300

QY 301 TCGGAGCGCTTGGGTGTTCGAGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 TCGGAGCGCTTGGGTGTTCGAGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTG 360

QY 361 TGGACTTACCGGAGCGCTGCTGCAATGACCGCGCTGAGCGGCTTACCTGGCCATC 420
Db 361 TGGACTTACCGGAGCGCTGCTGCAATGACCGCGCTGAGCGGCTTACCTGGCCATC 420

QY 421 TGGCGCGCGCTGCGGCGCGCTGTGTGATCACCGGCGCGCGCTGCGGCGCTATCGCT 480

Db 421 TGGCGCGCGCTGCGGCGCGCGCTGTGTGATCACCGGCGCGCGCTGCGGCGCTATCGCT 480

QY 481 GTGCTTGGGCGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GTGCTTGGGCGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 541 CAGGACCGCGGAGCTTCCGTATGCTGCGGAGCTTCAATGAGCACCGGCGGATGCGCTCC 600
Db 541 CAGGACCGCGGAGCTTCCGTATGCTGCGGAGCTTCAATGAGCACCGGCGGATGCGCTCC 600

QY 601 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 661 GGGCGCGAGACCGGAGGCGCGGCGCTGCTGAGCGCGGATGCGGCGGAGCGCGCG 720
Db 661 GGGCGCGAGACCGGAGGCGCGGCGCTGCTGAGCGCGGATGCGGCGGAGCGCGCG 720

QY 721 CAGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 CAGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 781 CTGTGCTGAGCATCTCTTACGAGGCTCATCGGCGGAGCTGTGAGAGCGCGCGCG 840
Db 781 CTGTGCTGAGCATCTCTTACGAGGCTCATCGGCGGAGCTGTGAGAGCGCGCGCG 840

QY 841 CTGCGAGGCGCGCGCGCGCTGCGGCGGAGAGAGCCACCGGAGCAACGCGCTCG 900
Db 841 CTGCGAGGCGCGCGCGCGCTGCGGCGGAGAGAGCCACCGGAGCAACGCGCTCG 900

QY 901 CGTAAGTGAAGCGCGCGCTGCTTCAAGAAGCGCTGAGTCCGCGCGCGCGGAGCG 960
Db 901 CGTAAGTGAAGCGCGCGCTGCTTCAAGAAGCGCTGAGTCCGCGCGCGCGGAGCG 960

QY 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 GCGCAACGCTGGGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 1021 TTTCCATTTTCATTCAGCTTCCACCGCGCG 1052
Db 1021 TTTCCATTTTCATTCAGCTTCCACCGCGCG 1052

RESULT 5
AX154589 1239 bp DNA linear PAT 22-JUN-2001

LOCUS
DEFINITION Sequence 9 from Patent WO0138355.
ACCESSION AX154589
VERSION AX154589.1 GI:14536175
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1239)
Sheppard, P.O., Jaspers, S.R., Deisher, T.A. and Bishop, P.D.
Method of forming a peptide-receptor complex with zsig33 and
therapeutic use thereof
Patent: WO 0138355-A 9 31-MAY-2001;
ZymoGenetics, Inc. (US)
FEATURES
Location/Qualifiers
1..1239
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1239
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC42611.1"
/db_xref="GI:14536175"
/translation="MGSPMNGSDGPEGAREPPWALPPEDRRCSPPFLGALVPTAV
PWFPLICRSLIVGSGCTATLLHMTALSVERYALICPLRARVLVTRRRVALIA

DB	440	TGCGGACCGCTCCGCGCGCCGCTTTTGTATCACC	CGGAGCGCGCTCCGCGGCTCATCGCT	499
OY	481	GTGCTCTGGGCGCGTGGCGGCTGCTCTGCGCGGCTCTCTGTTCTGTGTTGGGCGCTCGAG	540	
Db	500	GTGCTCTGGGCGCGTGGCGGCTGCTCTGCGCGGCTCTCTGTTCTGTGTTGGGCGCTCGAG	559	
OY	541	CAGGACCCCGGCGATCTCCGTAGTCCCGGCGCTTCATAGGACCGCGCGATCGCTCTCG	600	
Db	560	CAGGACCCCGGCGATCTTCGCTAGTCCCGGCGCTTCATAGGACCGCGCGATCGCTCTCG	619	
OY	601	CCCTGTGCGCTCTGCGCGCGCTCTCTGAGCTCTGCGGCGCGCCACCGCGCTCCCGCGCTCG	660	
Db	620	CTTTTGCTGTTGTTGCGCGCGCTCTTTTGGG-----TCTGCGGCGCGCACCGCGTCCCGCTCG	674	
OY	661	GGGCGCGGAGCGCGCG	675	
Db	675	GGGCGCGGAGCGCGCG	689	
RESULT 7	HSAs42408	692 bp	DNA	linear
LOCUS	HSAs42408			
DEFINITION	Human sapiens genomic sequence surrounding NotI site, clone			
ACCESSION	U01342.408			
KEYWORDS	U01342.408.1	GI:15886883		
SOURCE	Human sapiens			
ORGANISM	Human sapiens			
REFERENCE	Human sapiens			
AUTHORS	Kubenko, A.S., Gzatsullin, R.Z., Al-Amin, A.N., Wang, F., Kvasha, S.M., Podomov, I.R.M., Matushkin, Y.G., Gyanchandani, A., Muravenko, O.V., Levitsky, V.G., Kolchanov, N.A., Protodanov, A.I., Kashuba, V.I., Kisseliev, L.L., Wasserman, W., Wahlstedt, C. and Zabarovsky, F.R.			
TITLE	NotI flanking sequences: a tool for gene discovery and verification of the human genome			
JOURNAL	Nucleic Acids Res.	30 (14)	3163-3170	(2002)
MEDLINE	22131767			
PUBMED	12136098			
REFERENCE	2 (bases 1 to 692)			
AUTHORS	Zabarovsky, F.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-MAY-2001)			
FEATURES	Microbiology and Tumorbiology Centre, Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77, Sweden			
source	Location/Qualifiers			
	1..692			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="NRI-PB24C"			
BASE COUNT	92 a 239 c 219 g 140 t			2 others
ORIGIN				
Query Match	28.0%;	Score 572.2;	DB 9;	Length 692;
Best Local Similarity	.91.5%;	Pred. No. 8.1e-79;		
Matches	616;	Conservative 0;	Mismatches 55;	Indels 2;
				Gaps 1;
OY	1	ATGGGACAGCCCTTGGAACGCGACGCGAGCGCCCGAGGGGCGCGGAGCGCGCTGAGCCC	60	
Db	20	ATGGGACAGCCCTTGGAACGCGACGCGAGCGCCCGAGGGGCGCGGAGCGCGCTGAGCCC	79	
OY	61	GCGCTGCGCGCTTGGAACGCGAGCGCGCGCTGCGCCCTTCCCTGGGGGCGCTGTGCGG	120	
Db	80	GCGCTGCGCGCTTGGAACGCGAGCGCGCGCTGCGCCCTTCCCTGGGGGCGCTGTGCGG	139	
OY	121	GTGACCGCTGTGTGCGCTTGTGCTGTGCTGCGGGGTGAGCGGCAAGTGTGACCGTG	180	
Db	140	GTGACCGCTGTGTGCGCTTGTGCTGTGCTGTCGTCGAGGGGTGAGCGGCAAGTGTGACCGTG	199	
OY	181	ATGTCGATCGGGCGCTTACCGGAGATGCGGACACACACCACTTGTACTTGGGCGCATG	240	

Db	200	ATGCTATCGGGGGCTACCCGGACATCCGGACCAACCACTTGATCCTGGGCACATG	259
Qy	241	GCCGTGTCCGACCTACTCATCTCTGCTCGGAGCTGCCTGACCTGATACCCCTCTGAGCG	300
Db	260	GCCGTGTCCGACCTACTCATCTCTGCTCGGAGCTGCCTGACCTGATACCCCTCTGAGCG	319
Qy	301	TGCGGAGCCCTGGAGTGTTCGGGGCCGCGCTCTGCGCGCTCTCCCTCTAAGTGGCCAGGGC	360
Db	320	TGCGGAGCCCTGGAGTGTTCGGGGCCGAGATGCTATGCTGATCCCTTACGTGGGCGAGGGC	379
Qy	361	TGCACCTACGCCACAGCTGCTGACATGACCGGCTAGGCTGAGCGCTACTGAGCCATC	420
Db	380	TGCACCTACGCCACAGCTGCTGACATGACCGGCTAGGCTAAGGCTAAGCTTGGCATC	439
Qy	421	TGCGGAGCCCTCGGCGCCCGCGCTTTGGTTCACCCGGGCGCGGCTCATGCT	480
Db	440	TGCGGAGCCCTCGGCGCCCGCGCTTTGGTTCACCCGGGCGCGGCTCATGCT	499
Qy	481	GTGCTCTGGGCGCGTGGCGCTGCTCTCTGCGCGGCTCTTGTTCGTGGTGGGCGTCCGAG	540
Db	500	GTGCTCTGGGCGCGTGGCGCTGCTCTCTGCGCGGCTCTTGTTCGTGGTGGGCGTCCGAG	559
Qy	541	CAGACATCCCGGATCTCCGTAGTCCCGGGCTCAATAGGACCGCGCGATCTCTCTCG	600
Db	560	CATGACACCTGTATCTCCGTAGTCTCGGGCTCAATAGCACCGGCGGATAGCTACTTC	619
Qy	601	CGTCTCTGGGCGCGTGGCGCTGCTCTCTGCGCGGCTCTTGTTCGTGGTGGGCGTCCGAG	658
Db	620	GCTTATAGCATAGTGTGACGACTCTGTGTATTCGGGGGTGACACCATGATATACCT	679
Qy	659	CGGGGCGCCGAGAC	671
Db	680	CAGGCGCCGATAC	692
RESULT 8			
HSBJ39459			
LOCUS			
DEFINITION	615 bp	DNA	linear PRI 18-JUL-2002
ACCESSION	HSBJ39459		
VERSION	NRI-WC14C		
KEYWORDS	AJ339459		
ORGANISM	AJ339459.1	GI:15883877	
SOURCE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			

TITLE	JOURNAL	REFERENCE	AUTHORS
-------	---------	-----------	---------

Mus musculus, clone RP24-199E5

Anderson, S., Baerla, N., Baetien, V., Bogdanovskiy, L., Boukhalter, B., Brown, A., Camarata, J., Campolillo, A., Chang, J., Chazaro, B., Choedel, I., Colangelo, M., Coppolino, S., Collymore, A., Cook, A., Cooke, P., Dearlitzhu, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gadyas, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hager, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamarets, R., Landers, T., Lehoczkvy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menais, L., MiNova, T., Mlenga, V., Murphy, I., Naylor, J., Nguyen, C., Nicol, R., Nobbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Punnkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rossetti, M., Roy, A., Santos, R., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Spencer, B., Stange-Rothmann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talmes, J., TeeFay, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vasiliiev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaimont, S., Zember, L., Zimmer, A., and Zody, M.

[illegible]

submitted 03-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 172168)

Anderson, S., Barna, N., Baskett, C., Lander, E., Ali, A., Allen, N.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, J., Collins, S., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Guyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamati, A., Karatas, A., Kells, C., Labèque, K., Lamazates, R.,
Lancien, T., Macdonald, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Madersen, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, W., McEwan, P., McKernan, K., Meldrum, J., Meutus, L.,
Milnova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhking, P., Pierre, N., Pollars, V.,
Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rossetti, M., Roy, A., Santos, R., Schauer, S., Schpack, R.,
Seaman, S., Severyn, P., Spencer, B., Stange-Thomann, N., Stanjovic, N.,
Strassburg, N., Strzemiennian, A., Talamaia, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted 325-Apr-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 25, 2002 this sequence version replaced gi:18462076.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 Protocol: www-genome.washington.edu/mv/RepeatMasker.html

```
----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.960731
Consensus quality: 16510 bases at least Q40
Consensus quality: 167685 bases at least Q30
Consensus quality: 166765 bases at least Q20
```

Insert size: 176000; agarose-fp
Insert size: 169868; sum-of-contigs
Quality coverage: 8.4 in Q20 bases; agarose-fp
Quality coverage: 8.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.

- * this sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

1	291:	contig of 291 bp	in length
292	391:	gap of 100 bp	
392	1443:	contig of 1052 bp	in length
1444	1543:	gap of 100 bp	
1544	2462:	contig of 919 bp	in length
2463	2562:	gap of 100 bp	
2563	3236:	contig of 674 bp	in length
3237	3336:	gap of 100 bp	
3337	4351:	contig of 1017 bp	in length
4354	4451:	gap of 100 bp	
4454	5451:	contig of 1000 bp	in length
5454	5551:	gap of 100 bp	
5554	6971:	contig of 1418 bp	in length
6972	7071:	gap of 100 bp	
7072	8117:	contig of 1066 bp	in length
8118	8237:	gap of 100 bp	
8238	10559:	contig of 2332 bp	in length
10570	10669:	gap of 100 bp	
10670	13572:	contig of 2903 bp	in length
13573	13672:	gap of 100 bp	
13673	16655:	contig of 2983 bp	in length
16566	16755:	gap of 100 bp	
16756	20651:	contig of 3880 bp	in length
20636	20735:	gap of 100 bp	
20736	22907:	contig of 2172 bp	in length
22908	23007:	gap of 100 bp	
23008	26263:	contig of 3356 bp	in length
26264	26363:	gap of 100 bp	
26364	31037:	contig of 4674 bp	in length
31038	31137:	gap of 100 bp	
31138	38064:	contig of 6627 bp	in length
38065	38164:	gap of 100 bp	
38165	48247:	contig of 10083 bp	in length
48248	48347:	gap of 100 bp	
48348	57922:	contig of 9575 bp	in length
57923	58022:	gap of 100 bp	
58023	71658:	contig of 13606 bp	in length
71629	71728:	gap of 100 bp	
71729	90376:	contig of 18648 bp	in length
90377	90476:	gap of 100 bp	
90477	107535:	contig of 17059 bp	in length
107536	107635:	gap of 100 bp	
107636	131187:	contig of 23552 bp	in length
131188	13287:	gap of 100 bp	
131288	152394:	contig of 21107 bp	in length
152395	152494:	gap of 100 bp	
152495	172168:	contig of 19674 bp	in length

FEATURES

SOURCE

```
misc_feature      1.291
                 /note="assembly_fragment"
misc_feature      392.143
                 /note="assembly_fragment"
misc_feature      1544.2462
                 /note="assembly_fragment"
```


Query Match	15.0%;	Score 307;	DB 2;	Length 172168;
Best Local Similarity	70.2%;	Pred. No. 4.5e-38;		
Matches 530;	Conservative	0;	Mismatches 180;	Indels 45;
				Gaps 7

Db	35217	TCATTCGCTCGCTCTGGGTTGTGGCCGCTGTTCT-----TAGTTCCTCTTCTGTGG	35225
QY	533	GGCTGAGCAGAGACCCCGGACATCTCCGTAGTCCCGGCGCTCAATGAGACCGCGGANTCG	592
Db	35224	GGGTGCATCAAGATGTCCTCCCGGACCTCTGTGGAGACTGGATCAACATTGGGAATCA	35165
QY	593	CCTCTCGCCTCTGTGCTCTGTGCGCGCTCTCTGTGGCTCTGGCGGCGGCACCGCGTCC	652
Db	35164	ACCCCTCGCTCCGGGTCTGTGAGCTTCTCTTGGGTCTCGAGACACCAACGCTGTCC	35105
QY	653	CGCGTCGGGGCCCGAGACCGCGAGGCGCGGCGCTTTGACCGCGAATGCGGCGGA	712
Db	35104	CGCAATTGAGAACCAAGGCTCAGAGAGC-----GTTAGCGCGAGTGCC-ATGGA	35055
QY	713	GGCCCGCGCAGCTGGCGCGCGTGCCTGTATCTAGTGTGGGTACCAACCGCTACTTCTTCC	772
Db	35054	GGCCCACTCAGGTAGAGGCTCTGGGGGTCGTGTGTGGTACCAACCGCCCTA---TTTCC	34998
QY	773	TGCCCCCTTCTGTGCTCAGCATCTCTACAGGGCTCATGGGCGGGAGCTGTGGAGCACC	832
Db	34997	TGCCCCCTTCTGGGCTCTCGGTGTCTCTCGCGGGCTCTGTG-GGGCGTAGCCGTGCAGAGCC	34939
QY	833	GGCGGCGCGCTGGAGAGCCCGGCGGCTCGTGGGCGG	867
Db	34938	GGGTCCCGCACTAAGGTGGGAGAGAAAGGGGCGG	34904

[illegible]

AR182284
 LOCUS AR182284 283 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 4 from patent US 6337206.
 ACCESSION AR182284
 VERSION AR182284.1 GI:20225200
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 283)
 Tan, C. and Kolakowski, L.F., Jr.
 Nucleic acid encoding mouse galanin receptor (GALR2)
 JOURNAL Patent: US 6337206-A 4 08-JAN-2002;
 FEATURES
 Source
 1..283
 Location/Qualifiers
 BASE COUNT 27 a 116 c 84 g 56 t
 ORIGIN
 Query Match 13.9%; Score 283; DB 6; Length 283;
 Best Local Similarity 100.0%; Pred. No. 3.9e-34;
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 206 TGGGACCAACCACTGTAAGTGGGAGCATGCGCTGTCGACTTACTTCTGCTGC 265
 DB 1 TGGGACCAACCACTGTAAGTGGGAGCATGCGCTGTCGACTTACTTCTGCTGC 60
 QY 266 TCGGAGCTCCGTTCCAGCTGTACCGGCTGCGCTGCGGCGCTGGGTGTTGGGCGCG 325
 DB 61 TCGGAGCTCCGTTCCAGCTGTACCGGCTGCGCTGCGGCGCTGGGTGTTGGGCGCG 120
 QY 326 TGCCTGCGCGCTGTTCCTCTTACGTGGGCGAGGGCTGACCTTACGCGAGCTGTGACCA 385
 DB 121 TGCCTGCGCGCTGTTCCTCTTACGTGGGCGAGGGCTGACCTTACGCGAGCTGTGACCA 180
 QY 386 TGACCGGCTCAGCGGTGAGGCGCTACCTGGCCATGCGCGCGCTGCGCGCGCTGCT 445
 DB 181 TGACCGGCTCAGCGGTGAGGCGCTACCTGGCCATGCGCGCGCTGCGCGCGCTGCT 240
 QY 446 TGGTCAACCCGCGCGCGCTGCGCGCTGCGCGCTGCTGCTGCTG 488
 DB 241 TGGTCAACCCGCGCGCGCTGCGCGCTGCGCGCTGCTGCTGCTG 283
 RESULT 13
 HSA341295 573 bp DNA linear PRI 18-JUL-2002
 LOCUS HSA341295
 DEFINITION Homo sapiens genomic sequence surrounding Nct1 site, clone
 N5-KN11C.
 ACCESSION AJ341295
 VERSION AJ341295.1 GI:15885713
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 573)
 Kutsenko, A.S., Gizaullin, R.Z., Al-Amin, A.N., Wang, F., Kvasna, S.M.,
 Podowski, R.M., Matuchkin, Y.G., Gyanchandani, A., Muravenko, O.V.,
 Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I.,
 Kisselev, L.L., Wasserman, W., Wahlestedt, C., and Zabarovsky, E.R.
 Nct1 flanking sequences: a tool for gene discovery and verification
 of the human genome
 JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)
 MEDLINE 22131767
 REFERENCE
 2 (bases 1 to 573)
 Zabarovsky, E.R.
 TITLE Direct Submission
 AUTHORS Submitted (16-MAY-2001) Microbiology and Tumorigenology Centre,
 Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
 Sweden
 FEATURES
 Location/Qualifiers

source 1..573
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="N5-KN11C"
 BASE COUNT 64 a 185 c 195 g 129 t
 ORIGIN
 Query Match 13.1%; Score 266.8; DB 9; Length 573;
 Best Local Similarity 98.9%; Pred. No. 1.2e-31;
 Matches 279; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 ATGGGAGCCCTTGGAAAGCGACGACGCGCCCGAGGAGGCGGAGCGCGCTGGGCC 60
 DB 20 ATGGGAGCCCTTGGAAAGCGACGACGCGCCCGAGGAGGCGGAGCGCGCTGGGCC 79
 QY 61 GGGCTCCGCTTGGAGAGAGCGCGCGCTGCTGCGCCCTTCCCTGGGGCGCTGGTCCG 120
 DB 80 GGGCTCCGCTTGGAGAGAGCGCGCGCTGCTGCGCCCTTCCCTGGGGCGCTGGTCCG 139
 QY 121 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 140 GTGACCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 199
 QY 181 ATGCTGATGGGCGCTTACCGGAGATGCGGACCAACCACTTGTACTTGGGAGCATG 240
 DB 200 ATGCTGATGGGCGCTTACCGGAGATGCGGACCAACCACTTGTACTTGGGAGCATG 259
 QY 241 GCGGTGTCGACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 282
 DB 260 GCGGTGTCGACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 RESULT 14
 AR156351 1063 bp DNA linear PAT 08-AUG-2001
 LOCUS AR156351
 DEFINITION Sequence 1 from patent US 6242199.
 ACCESSION AR156351
 VERSION AR156351.1 GI:15125055
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 1063)
 Ploeg, L.H.T., Feigheuer, S.D., Howard, A.D., Pong, S.-S. and Van Der
 TITLE Assays for growth hormone secretagogue receptors
 JOURNAL Patent: US 6242199-A 1 05-JUN-2001;
 FEATURES
 Source
 1..1063
 Location/Qualifiers
 BASE COUNT 166 a 339 c 285 g 253 t
 ORIGIN
 Query Match 11.5%; Score 234.8; DB 6; Length 1063;
 Best Local Similarity 69.3%; Pred. No. 9.6e-27;
 Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
 QY 92 GCGCTTCCCTTCCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 151
 DB 69 GCGCTTCCCTTCCGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128
 QY 152 TCGGAGTGGCGGAGCAAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 211
 DB 129 TCGGAGTGGCGGAGCAAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 188
 QY 212 CCACCAACCACTTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 271
 DB 189 CCACCAACCACTTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 248
 QY 272 TGGCGTTGACCTGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 331
 DB 249 TGGCGTTGACCTGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308

QY 332 GCGCCTGTCCTCTACGTGGGCGAGGCTGCACCTACGCGACCGCTGTCGACATGACCG 391
 DB 309 GCAACTCTTCCAGTGTGTAGGAGAGCTGCACCTACGCGACAGTGTCTACCTACCG 368
 QY 392 CGCTAGCGCTGAGCGCTACCTGCGCATCTGCGCGCGCTCGCGCGCGCTTGTGCA 451
 DB 369 CGCTAGCGCTGAGCGCTACCTGCGCATCTGCGCGCGCTCGCGCGCGCTTGTGCA 428
 QY 452 CCGCGCGCGCGCTCGCGCGCTGCTGCTGCTGCGCGCGCTGCTGCTGCTGCGCG 511
 DB 429 CCAAGGCGCGCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCG 488
 QY 512 GTCCTCTTCTGTCCTGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCTGCTGCGCG 553
 DB 489 GGCCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCG 530

RESULT 15
 SSU60178
 LOCUS
 DEFINITION Sus scrofa growth hormone secretagogue receptor type 1a mRNA,
 complete cds.
 U60178
 U60178.1 GI:1504144

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Sus scrofa.
 Sus scrofa.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 1101)
 Howard, A.D., Feighner, S.D., Cully, D.F., Arena, J.P., Liberato, P.A.,
 Rosenblum, C.I., Hamelin, M., Hreniuk, D.L., Palyna, O.C., Anderson, J.,
 Pares, P.S., Diaz, C., Chou, M., Liu, K.K., McKee, K.K., Pong, S.-S.,
 Chaung, L.-Y., Eibrecht, A., Dashkevich, M., Heavens, R., Rigby, M.,
 Strimachsingh, J.D.J.S., Dean, D.C., Melillo, D.G., Patchett, A.A.,
 Nargund, R., Griffin, P.R., Demartino, J.A., Gupta, S.K.,
 Schaeffer, J.M., Smith, R.G. and Van der Ploeg, L.H.T.
 A receptor in pituitary and hypothalamus that functions in growth
 hormone release
 Science 273 (5277), 974-977 (1996)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 Submitted (06-JUN-1996) Genetics & Molecular Biology, Merck
 Research Laboratories, P.O. Box 2000, Building RY-80Y-265, Rahway,
 NJ 07065, USA

FEATURES
 Source
 CDS
 Location/Qualifiers
 1..1101
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /tissue_type="anterior pituitary"
 1..1101
 /function="growth hormone release"
 /note="G-protein-coupled receptor of pituitary and
 hypothalamus; receptor for growth hormone releasing
 peptides (e.g. GHRP-6, GHRP-2) and non-peptide, low
 molecular weight secretagogues (e.g. L-692,429, MK-0677)"
 /codon_start=1
 /product="growth hormone secretagogue receptor type 1a"
 /protein_id="AAC48630.1"
 /db_xref="GI:1504145"
 /translation="WMNATPSEBPGNLTIPIGDAPENDSTVBEILPLPPTPLA
 GVTATCAALFVAGLNGNLITMLVSRFRMRTTNLYSSMAFSLILFLCPDLFR
 LMQYPMNLGNLKLQFVSESCYATVLTALSVRFPAICFPLRAKAVITGRV
 KLIVLIVAVAFCSAGPIFVLGVGEHDNGTDPDNECRATEFAVRSGILITVMWVSS

VPEFLPVLCITLVYSLIGKILRRKREAAVSSLRDQNHQYKMLAVVPAFLICM
 LPRHVGRIYFSKSLPESVETIAQISQYCNVSVFLVFLYSAINPILNIMSKINAV
 FKLIGFPPFQRLKSLTKDESSRAMTSSINT"
 BASE COUNT 196 a 351 c 299 g 255 t
 ORIGIN
 Query Match 11.5%; Score 234.8; DB 4; Length 1101;
 Best Local Similarity 69.3%; Pred. No. 9.5e-27;
 Matches 320; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 92 GCGCCTTCCCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 151
 DB 107 CGCTTCCCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 166
 QY 152 TCGGAGTGAAGCGCAAGCTGTGAGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 211
 DB 167 TGGATATGCGCGCGCAAGCTGTGAGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 226
 QY 212 CCACCAACAATTGTAAGCTGTGAGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 271
 DB 227 CCACCAACAATTGTAAGCTGTGAGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 286
 QY 272 TGGCGTTGACCTGTACCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 331
 DB 287 TGGCGTTGACCTGTACCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 346
 QY 332 GCGCCTGTCCTCTACGTGGGCGAGGCTGCACCTACGCGACCGCTGTCGACATGACCG 391
 DB 347 GCAACTCTTCCAGTGTGTAGGAGAGCTGCACCTACGCGACAGTGTCTACCTACCG 406
 QY 392 CGCTAGCGCTGAGCGCTACCTGCGCATCTGCGCGCGCTCGCGCGCGCTTGTGCA 451
 DB 407 CGCTAGCGCTGAGCGCTACCTGCGCATCTGCGCGCGCTCGCGCGCGCTTGTGCA 466
 QY 452 CCGCGCGCGCGCTCGCGCGCTGCTGCTGCTGCGCGCGCTGCTGCTGCTGCGCGCG 511
 DB 467 CCAAGGCGCGCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCGCG 526
 QY 512 GTCCTCTTCTGTCCTGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCTGCTGCTGCGCG 553
 DB 527 GGCCTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCG 568

Search completed: March 16, 2003, 03:14:23
 Job time : 6757 secs